DOCUMENT RESUME

ED 350 694 EA 024 383

AUTHOR	Wendel, Frederick C.; And Others
TITLE	Measurements of Personality and Leadership: Some Relationships.
PUB DATE	[92]
NOTE	121p.
PUB TYPE	Viewpoints (Opinion/Position Papers, Essays, etc.) (120)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Assessment Centers (Personnel); Elementary Secondary Education; *Leadership; *Leadership Qualities; Measurement Techniques; Personality; *Personality Measures; *Testing Problems

ABSTRACT

The relationship between measurements of personality and leadership is described in this paper. Following an introduction and a statement of the problem, the first part discusses the principles of measurement, beginning with the basic issues and assumptions and going on to examine various aspects of validity, reliability, instrument construction, and performance testing. The second part discusses personality and its definition, personality and leadership, traits and criticism of traits, adult personality development, identity, and personality testing. Issues in the study and measurement of leadership are described in the next part, with a focus on the uses of different instruments to measure leadership. Methods used by assessment centers are examined in the fourth part, with attention to the pros and cons of psychological testing and leaderless group discussions. The final part summarizes topics covered in each of the preceding sections. One conclusion is that the assessment center method is effective for direct observation of job-related skills, but that tests should be carefully selected, administered, and scored. The design of personnel policy must correspond with the use of appropriate measures. (Contains 86 references.) (LMI)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED350694

Measurements of Personality and Leadership: Some Relationships

Frederick C. Wendel
Professor of Educational Administration
1204 Seaton Hall
University of Nebraska-Lincoln
Lincoln, Nebraska 68588-0638
(402) 472-3726

Allan H. Schmidt
Concordia College
Seward, Nebraska

James Loch
Hastings, Nebraska

BEST COPY AVAILABLE

Running head: MEASUREMENTS RELATIONSHIPS

EA 024 383

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
F.C. Wendel
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

2

ERIC

TABLE OF CONTENTS

Introduction

Principles of Measurement

Personality

ERIC

Leadership

Assessment Center Methodology

Summary and Conclusions

Measurements of Personality and Leadership: Some Relationships

Introduction

Background

People have intense interest in measuring, testing, comparing, and obtaining determinants of quantity or quality. Whether the object is animate or inanimate is often of little consequence to individuals who are measuring or comparing.

Values of measurement. What nation holds which records in space exploration is a matter of great concern to leaders of state. Every mother has the "absolute cutest" baby. In sports, the won-lost record at the end of the season determines the winner--and the losers. One person emerges from a job applicant pool as the new hire; although the new hire may be overjoyed at being selected, the elation of the employer may turn sour if the newly hired employee does not measure up to expectations.

How do government officials determine which corporation of scientists, engineers, mathematicians, financial experts, managers, etc. is going to design and manufacture the best rocket to put the biggest payload into space first? Which baby can be the "cutest" when most are look alikes for Winston Churchill? Who can tell before a season begins who will win the Davis,

ERIC

Lombardi, or some other emblematic cup? How can hiring officials determine which applicant is the best person for a job?

Measurement of job success. One follow-up study, 10,000 Careers, produced some 12,000 correlations of test scores with indicators of job success (Thorndike, 1963). During World War II, some 75,000 applicants for aircrew training were given a battery of tests that yielded 20 separate scores on verbal, numerical, spatial, perceptual, and motor abilities. The battery was used to assign men to officer training for pilot, navigator, or bombardier or to enlisted status as a tail gunner, radio operator, or other member of a bomber crew. About 10 years after the war, 17,000 names were randomly selected for study, and records on 10,000 men were available. These men had assumed positions from corporate board rooms to prison cells. What were the findings?

> I think the simplest and most honest way of summarizing these results would be to say that they clustered around zero--almost as many negative as positive--with the number "significant at the five per cent level" making up perhaps six percent of the whole. Of these presumable "significant" correlations, about as many were in the "silly" direction as in the sensible one. Applying these results, taken at face value to the question, "Can tests given at about age 20 predict occupational success 12 years later?" the answer is clearly, "No." (Thorndike, 1963, 182).

ERIC

<u>Problems with the measurement of job success</u>. Nine propositions for the tests' failure to predict occupational success included (Thorndike, 1963): (a) the group was too homogeneous, (b) the tests were unrelated to civilian jobs, (c) the tests were selected for training purposes and not for predicting job success, (d) the tests were not designed to differentiate among occupations, (e) the tests covered a limited range of abilities and missed important ones, (f) the abilities that were tested varied over time, (g) the men after the war were in occupational groups so diverse that no tests could distinguish occupational success, (h) occupational success (two chief measures were salary level and promotions) was not adequately evaluated, and (i) occupational success is a function of so many unknown contingency factors that prediction is impossible.

Although the lack of correlations with job success from <u>10,000 Careers</u> was disappointing, Thorndike ascribed "Some part of the reason for such negative results may lie in our somewhat curtailed group, and more of it may be found in the incomplete coverage of possible significant domains of interest, temperament, and ability" (1984, p. 186). Thorndike also drew some valuable conclusions: (a) there is great heterogeneity among occupations, (b) there are many differences between training and work, (c) pay and promotions may be so institutionalized that there is little

meaning with job success, and (d) contingency factors may weigh heavily upon careers.

Importance of personal selection. Hiring officials, nevertheless, want the best means of selecting the right individual for a particular job. Businesses, colleges, government agencies, schools, and other employers spend thousands of dollars in direct and indirect costs in figuring out what a specific job's requirements are and what changes, if any, should be made in them; forecasting personnel needs; recruiting qualified applicants, internally and externally; selecting among qualified applicants; and evaluating the selection process.

The selection process, if conducted thoroughly, is expensive. It requires expenditures for preparing, printing, distributing, collecting, and analyzing application forms. Paper-and-pencil questionnaires and tests may need to be designed, selected, collected, and analyzed. One or more rounds of interviews with one or more individuals, teams, or groups must be designed, scheduled, conducted, and analyzed. Reimbursing applicants' expenses, at least for top-level jobs, may be another cost. Weighing each applicant's abilities against job requirements and comparing all applicants' abilities against each other takes hiring officials' time. Scheduling and conducting meetings of hiring officials to confer, evaluate, discuss, and debate the

ERIC

perceptions about applicants may consume valuable resources. Offering the top applicant the job and thanking the other applicants for their interest in the job has its costs. Filing and storing applications and reports to comply with Equal Employment Opportunity Commission guidelines add to the costs. In the worst case scenarios, the process may need to be duplicated because the top choice has decided not to take the job or none of the applicants seems to fit the vacancy.

With the hundreds, and more typically, thousands of dollars spent on hiring a new employee, the need for good selection processes and procedures is evident. Even the readily direct expenditures of a selection process can pale in comparison with the failure to select a successful employee. Job turnover is expensive even at the lowest level, but at management levels, the impact of poor hiring decisions is vastly more costly than outlays for hiring officials' time in planning, interviewing, evaluating applications, and the like. What is the cost if a leader heads in a wrong direction? The cost of paying an ineffective administrator thousands of dollars a year for one or two years is readily apparent but putting a price tag on lost opportunities is nearly impossible.

Past, present and future. In the corporate world, the cost of poor management has a bottom line (Kotter, 1988).

ERIC

> It was a whole lot easier to be an executive thirty years ago. Back then, there were lots of opportunities for growth. Today, there is more competition and our markets are much more mature. When I first joined the company in 1952, we actually had monthly "allocation meetings" in our division, meetings in which we decided which customer got our products. Can you believe that?
>
> Today, we need many more and better leaders than back then, broad people with vision and self-confidence. Without these people, there is no way we will continue to prosper. In some of our businesses, without them we won't even survive (Kotter, 1988, p. 10).

How would school executives echo the corporate officer? They might say something like this: "It was a whole lot easier to be a school administrator thirty years ago. Back then, there was lots of growth in the community. Today, there is more competition for every kid's attention, more family problems, drug problems that hardly existed then, more need to contain costs, and more criticism of education. When I first joined the district in 1952, we actually had monthly 'allocation meetings' in our school, meetings in which we decided which teachers and kids got to transfer into the better programs. Can you believe that? Today, we need many more and better leaders than back then, broad people

with vision and self-confidence. Without these people, there is no way we can meet kids' needs. In some of our schools, without them we won't even survive if ideas like voucher systems catch on with the public."

Corporate and school leadership. A comparison with a corporation might provide further insights into what businesses and school executives face in the 1990s.

> At the same time that increased competitive intensity has been producing the need for more leadership at almost all levels in many organizations, a second set of less dramatic forces has been steadily increasing the difficulty of providing effective leadership. They are the forces of growth, diversification, globalization, and technological development, which have been making businesses more and more complex. PepsiCo is not unusual in that regard. In 1955, the Pepsi Cola Company was a $60-million-a-year soft drink firm that sold its product mostly in the United States and employed around 1,900 people. Twenty-five years later, it was a $6-billion-a-year corporation, with more than 100,000 employees, that sold soft drinks, snack foods (Frito-Lay), fast food (Pizza Hut and Taco Bell), transportation services (North American Van Lines and Lee Way Motor Freight) and sporting goods (Wilson) and that derived a significant amount

ERIC

of money from more than a hundred markets outside the United States. Providing effective leadership to the business in 1955 was probably not easy by any reasonable measure. But providing effective leadership in 1980 was definitely more complex by orders of magnitude (Kotter, 1988, p. 12).

In the last 25 years what changes have taken place in schools that rival those in corporations like PepsiCo? Is there a call for "increased competitive intensity" for instruction in science, math, foreign languages, and other subjects with Japan, Germany, and other nations? Is there a "need for more leadership at almost all levels in many schools?" Although the Chicago School Reform was struck down, the intent of the Illinois legislature to make schools more responsive to local needs highlights efforts to fill a leadership vacuum. Are there "forces of growth, diversification, globalization, and technological development which have been making schools more and more complex?" Think of growth in school-age population, the percentage of children with special needs, knowledge, research findings, mandated services to students, and expectations for school services. Think of diversification of learners' needs, learning styles, organizational plans to combat segregation, and staffing patterns. Think of the effects of globalization upon American education with comparisons of the length of the school day, week,

and year to test scores in science, math, and the like. Think of the technological developments which challenge schools. With limited budgets, school administrators must peer into the future to select the right computer system to network, process student and financial records, meet instruction and instructional management needs, and other elements that have yet to be invented.

Facing challenges. How well can school executives relate to the challenges that face business leaders?

> Dealing with the typical leadership challenges created by competitive intensity--getting costs down, increasing productivity, improving customer service, keeping quality high, getting new products developed faster--is rarely easy. Dealing with those issues always means producing change. Change creates uncertainty, anxiety, winners, and losers. The resistance generated by anxious people or employees facing real issues is seldom easy to overcome, even in simple situations. But simple is not the order of the day any more. And dealing with those challenges in complicated settings can be enormously difficult (Kotter, 1988, p. 13).

To what extent do school administrators have to be concerned with costs, productivity, service, quality, change, resistance, and complexity? But are all the problems external to executives?

ERIC

Do the pressures and problems come from "them" only? One observer of corporate culture has criticized the way leaders act.

> The culture of the American corporation continues to act on the assumed value of command leadership. Managers place a high value on the ability of the manager to make decisions himself, quickly. Allan Cox conducted a comprehensive survey of middle managers and executives in a sample group of U.S. corporations. He concluded that "approximately 70 percent of executives report that their corporations strongly or somewhat encourage both speedy decision making and a high energy level and fast pace on the job." Seventy-three percent of top executives reported that their companies valued quick decision-making ability; and 67 percent of middle-level managers reported the same value (Miller, 1984, p. 50).

Profiles of decision makers. How successful is the "Lone Ranger" type? How fruitful is one person's decision? Necessary at times, Yes, if one smells smoke and sees flames, the person won't wait to call a committee meeting to notify the fire department. Decisions can be made in three ways:

> Command decisions are those the individual manager makes without discussion with his subordinates or team members. Consultative decisions are those the manager makes but only after discussion, either one-on-one or in a group, with other

managers or persons who have knowledge or interests related to the decisions. Consensus decisions are those the manager turns over to a group of his peers or subordinates, and which are made after full and frank deliberation by the group and to which all agree to adhere as if the decisions were their own (Miller, 1984, p. 52).

Are there ways to determine how individuals make decisions, value the decisions of others, help others reach decisions, support leadership in others, seek leadership, and, in sum, provide effective leadership? Are there ways that improve the selection process of leaders? Since Thorndike's study of *10,000 Careers*, what has been learned about predicting job success? For example, are the domains of interest, temperament, and ability possibly significant in the identification of successful leaders?

Statement of the Problem

The identification of leaders is not an exact science. A decision to invest tens of thousands of dollars in the potential success of a corporate leader can invoke anxiety in the most self-confident chief executive officer or chairman of the board. Entrusting the education of thousands of young people to a superintendent, principal, or instructional leader is no less significant. The purpose of this monograph is to describe the relationship between measurements of personality and leadership.

Principles of Measurement

Basic Issues

Because a property, object, or abstraction may be so minute or huge, slow or fast, or difficult to identify, precise measurements can be made only with great difficulty. Specialists use carefully crafted instruments to measure characteristics of atoms or of the universe. The movement of glaciers poses different measurement problems than that of galaxies in our expanding universe. How do art experts decide whose paintings are more beautiful, Tintoretto's or Titian's, or which one by either artist is better than another?

Assumption of attribute. Scientists can measure some properties much more readily than others although exactitude to the billionth decimal place is but one step toward a more precise measurement. Scales, stop watches, and other instruments measure the height and weight of humans, clock their speed in races, and gauge blood pressure. Some persons make the "Ten Best Dressed" list because some expert has so determined. But beauty still remains in the eye of the beholder. Even a physical scientist's precise measurements of weight, height, and other attributes of objects are not necessarily those of a real thing (Dingle, 191). Because proof cannot be established that an underlying trait has been accurately identified, tested, and measured, judgments about

attributes of humans are fraught with imprecision and must be carefully drawn.

To add to the difficulty of measuring human attributes, consider intra-individual variability. Some objects may weigh the same day after day, but an attribute of an individual, i.e., independence, may not test out similarily. A study of management potential conducted by Standard Oil of New Jersey showed a cross-validated multiple correlation of .70 for identifying more effective and less effective managers (Dunnette, 1964).

> However, the managers did not fit a common mold when the various individual tests of the battery were examined. Some managers were high on one kind of test (such as verbal fluency); other managers, though low on this test, seemed to "make-up for it" by having a pattern of education and work experience in their backgrounds which compensated in a sense for their lowered measured ability (p. 64).

Also, consider inter-individual variability. To what exactitude can the trait of independence or mental ability be compared among several individuals or of people from different cultures?

Assumption of proof. Another assumption is that proof of something can be established by giving a test. At the end of the school year, achievement tests are commonly given. If the scores are high, there is "proof" that students are learning.

If the scores are not so high, "proof" of something else is established.

<u>Assumption of objectivity</u>. A standardized, objective test or inventory is not standard and objective in some absolute sense. The construction of any item, format of the test or inventory itself, wording of any phrase and item, and testing procedures all impose limitations upon standardization and objectivity.

<u>Assumption of precision</u>. A foot ruler is a handy device for measuring some objects if not too much precision is required, however, astro-physicists might prefer a different unit--the light year--in their measurements of galactic space. A test of mental ability, achievement, personality, or other attributes cannot be selected as one can get the candy bar of choice from a vending machine. A test may or may not measure precisely what it is designed to measure. While scientists may speak of "fundamental measurement," educators are more limited in measuring attributes for which there is a zero point to a broader means of grouping subjects into categories and ordering them (Thorndike, 1964).

Of equal importance is the role of measurement in making predictions of behavior. To what extent do navigators, cattle feed lot operators, stock brokers--all persons--rely upon past measures for future predictions? For the airline navigator, factors of miles to be traveled, wind speed, and aircraft

speed all have some dependable relationship. What increases the chances of success are good measures from past performance. Unfortunately, "The first, the most fundamental and the most disheartening fact is that all predictions are fallible. All have a component of error" (Thorndike, 1964, p. 105).

Predictions are generally less accurate in relationship to the time span over which they are extended. Predicting intellectual level from tests administered to preschoolers is far more fallible than predicting the same for adults.

Predictions also benefit from a close relationship between what is being predicted and the predictor. A reading test is a better means of predicting sixth graders' reading ability two or three years in the future than a test of verbal ability, of some other measure of intelligence, or of learning style. A better prediction of kindergarteners reading ability may come, however, from an individually administered intelligence test than from a reading test because of the limited reading ability of most kindergarteners. Predictions are enhanced by using available information that has the most obvious relationship to the factor.

Multitude of factors. Because of intra-individual variability, i.e., all behavior is a function of multiple determinants, predictions can be enhanced by using multiple factors that contribute to prediction (Thorndike, 1964). The use of multiple

factors, of course, raises other questions such as weighting and overlapping of the measures. Furthermore, judgments can be made more readily about status (present conditions) than about growth (future status). Likewise, judgments about the status of a single dimension can be made more readily than about differences between several dimensions.

Validity

One of the important features of measurement is validity. Validity indicates the degree to which a test or instrument achieves expressed aims. Tests are used for several purposes, and for each type, a different type of investigation must be used to establish validity.

There is no minimum validity coefficient that is used to separate acceptable from unacceptable tests. In terms of generally accepted experimental research statistics, validity should be above .50. Statistics for test instruments evaluating personality, leadership, and other abstract characteristics cannot always be judged by the same criteria as other types of research. A .50 validity coefficient on a personality test might be quite high, while in other areas of research a validity coefficient of .50 would be considered moderate to low. Any test instrument is never totally valid or invalid.

<u>Purposes of testing</u>. A common purpose of testing is to determine how an individual currently performs. An achievement test is designed to measure students' performance on a sample of questions that purport to represent a level of achievement.

Another purpose is to estimate a person's status on a variable or to forecast an individual's status in the future. For the former, a short inventory could estimate what the outcome might be from a full psychological examination. An academic aptitude test could be administered to forecast an individual's grades.

A third purpose of testing is to find out how much of some trait or quality a person has, to study the test itself, to develop a theory about a trait or quality, or to compare the test with that of another on a like quality. A mental ability test provides an indication of "intelligence" while another test might be used to infer the degree of creativity, dominance, or some other quality in a person.

Because tests are used for different purposes, test information must be gathered to determine the suitability of a test for its intended purpose. Validity information is thus related to the purpose of testing rather than to the type of test to be administered.

There are three types of validity (a) content validity, (b) construct validity, and (c) criterion-related validity.

Content validity. A measure of how well the content of a test samples the attribute, quality, subject matter or trait is known as content validity. Tests of achievement, adjustment, proficiency, and social behavior, for example, should demonstrate good content validity.

Construct validity. To what degree a concept accounts for performance on a test is a function of construct validity. Construct validity is a measure of the theory underlying a test.

Criterion-related validity. How well test scores compare with one or several external variables is shown by criterion-related validity. For example, expectancy tables and correlations of test scores with a criterion measure, e.g., graduation rate or job performance, are developed to show this information.

While content validity is important for measuring the factors under investigation and criterion-related validity is desirable for purposes of comparison, construct validity is perhaps the most significant and important kind of validity. Kerlinger (1979) defines it as "the psychological property or properties the instrument measures" (p. 139). This type of validity determines whether the actual theoretical bases of the information or skills to be measured are present in the test items.

Research is traditionally based upon a previously tested or a hypothesized theoretical base. Some of the most popular

personality and leadership instruments are not founded on a theoretical base, but on a hypothesized theory generated with questionable support from the literature. In contrast, many instruments are based on relatively complex theories and hypotheses. When a theory is complex, the information may not be clear and concise; and information received may present problems in interpretation and understanding. Because of the complexity and theoretical nature of this kind of validity, it can best be tested by the use of the instrument to investigate the constructs it was designed to measure. Once the instrument has been tested, construct validity becomes a matter of interpretation regarding what has been measured.

A personality inventory, for example, should demonstrate (a) construct validity, i.e., show a relationship with personality theory, and (b) criterion-related validity, i.e., show success in screening out maladjusted individuals.

Reliability

The stability and consistency of measurement by a test or inventory are referred to as reliability. At least two measurements must be available for comparison purposes. "Stability" describes performance over time while "consistency" describes performance over various forms of a test or inventory. Reliability is a measure of correlation. It often is indicated by

the letter "r" and a coefficient; therefore, it can range from + 1.00 to - 1.00.

Consistency of test scores may be affected by variations of responses by the test taker because of changes in mood, effort, or other psychological or physiological factors as fatigue. These factors may have great effect upon personality inventories.

Situational tests can be affected by the interactions of individuals. The variability of interactions of an individual may be relatively large from group to group as the groups are composed of different members.

Variability of responses may be due to administration or interpretation of the test. The skill of a rater or of a test administrator may produce inconsistent scores.

Variations in the process of observation can contribute to inconsistency of scores. Scoring-error variation can also be produced by errors in the scoring, recording, transferring, and reading of scores.

There are four major types of reliability that are frequently reported: (a) parallel or equivalent forms, (b) test-retest, (c) split-half, and (d) coefficient alpha. In parallel or equivalent forms, two tests containing items which are as equivalent as possible are correlated with each other. In the test-retest method of determining reliability, the identical

instrument is administered twice to the same individuals and the scores are then correlated. Both parallel or equivalent forms and test-retest are examples of external consistency of a test instrument.

The split-half method of reliability is based on one test administration only. By random arrangement, the scores of each person on half the test items are correlated with the person's score on the other half. Cronbach's coefficient alpha is the basic formula for determining reliability based on internal consistency. The coefficient indicates the average of all split halves, which includes all possible parameters. Typically, coefficient alpha is reported as the reliability score.

The retest method can be affected by response variability of subjects and by conditions of administrations of the test. A single administration, e.g., the "split-half method," of a test ignores response variability of the subject and administration conditions. Thus, any "reliability coefficient" must be accompanied by a description of the method used to derive the coefficient.

Clearly labeled components of error should be made known to test users so that they can interpret test scores properly. The correlation of an internal-consistency coefficient has a much

ERIC

different meaning from that of a correlation between two forms of a test administered on different days.

Instrument Construction

In determining the worth of any instrument there are a number of factors which need to be considered: (a) Purpose of the Instrument, (b) Constructs, (c) Norms, (d) Validity, (e) Reliability, and (f) Interpretation of Results.

Purpose of the instrument. A test user must know the purpose of the instrument, whether it will measure the particular constructs of interest to the researcher, and how accurately it will measure those constructs. Other important aspects include the number of references for the instrument, information from reviews that have been written, possible expectations from the test administration, and the populations for which the test is intended. The method(s) of instrument construction needs to be examined to understand the capabilities of the instrument and how it can be expected to perform. How the constructs were formulated and how scales were designed should be scrutinized closely. Not only are these aspects important for an understanding of how the test instrument may perform, but reliability and validity are closely related to their construction. Often too little time is spent in determining the purpose of an instrument.

Constructs. Constructs are understood to be the clusters of factors or categories to be evaluated by the instrument and are usually based on a research theory or one supported by the literature. The constructs may be based on a theory of a researcher, other test instruments, or intuition. The validity of the constructs is measured by how well they support the theory or hypothesis. Frequently, there are main headings (general) as well as a number of subcategories (specific), with scales used to interpret the results.

Norms. In determining the applicability of a test, the demographics of the population tested to establish the norms are an important consideration. Such factors as age, sex, grade, race, geographic region, and the normality of the population used are specific areas of interest. If the norms can be projected, to what populations can the norms be generalized?

Personality tests are frequently normed from clinical populations and may not be applicable to normal populations. Early leadership test instruments were normed with military personnel and may be of questionable value with certain populations.

Interpretation of results. How tests are scored is a matter of much debate, because certain items are often weighted. With rather illusive and vague fields of inquiry, such as personality

and leadership, the problem is compounded. Because of the difficulty in determining suitable scales, the data obtained from personality and leadership test instruments must be evaluated and interpreted carefully. Although many tests include a manual(s) to assist in understanding the results, information received may need to be examined and interpreted by a trained clinician or test administrator.

The results of leadership or personality tests must be evaluated with the purpose of the test as the focal point. Tests may be either predictive or descriptive, or a combination. A test may predict to what extent behaviors, styles, or traits are likely to occur or may provide a "snapshot" of factors as they are observed or measured at the time of testing. Instruments may combine the characteristics of different relationships to meet the purpose of that particular test.

Performance Testing

Performance testing is generally used to refer to non-verbal tests that require some form of behavioral response (Cronbach, 1970). Performance tests are more complex than paper-and-pencil tests and may require more complex behavioral responses. Responses on a performance test may vary from simply responding to a written or spoken word to working with a team of

individuals in a management game in competition against other teams or in some other complex task.

A situational test is a complex performance test that duplicates a realistic setting. Overt behaviors are usually evoked in situational tests; they generally do not seek a single answer.

Many early psychological testing programs consisted of performance tests that required subjects to participate actively by answering questions, memorizing material, or conducting psychomotor tasks. Much of these early tests, e.g., those conducted by Galton and Cattell, focused upon life- and work-related variables. These testing programs were of importance, not so much because of their findings, but because they began the practice of objective measurement of overt behaviors although problems of observation, scoring, rating, and other processes were not resolved.

The measurement of managers' behaviors was significantly enhanced by the development of the in-basket exercise (Thornton & Byham, 1982). This exercise was designed in the 1950s to measure officers' ability at the Air College to arrive at decisions, to organize information, to discover the problem in a complex situation, and similar administrative abilities. The success of the in-basket exercise brought about increased

efforts to develop other situational tests to measure managerial abilities.

Personality

The Study of Personality

What is personality? Can it be measured? What is its relationship to leadership? Do infants start life without any personality and develop one as they grow? Does personality change throughout life, or does a basic personality remain constant even though behavior seems to vary?

Definition. While there are many questions about personality, everyone knows what personality is. Personality is those characteristics that make an individual carefree and happy, withdrawn and secluded, or antisocial and violent. There are almost as many different definitions of personality as there are words to analyze and discuss it. According to Menninger (1953), personality:

> . . . has been used to describe almost anything from the attributes of the soul to those of a new talcum powder. As I shall use it, it means the individual as a whole, his height and weight and loves and hates and blood pressure and reflexes; his smile and hope and bowed legs and enlarged tonsils. It means all that anyone is and all that he is trying to become (p. 23).

Although Menninger's definition appears to be inclusive, he does not consider the view of the person that is taken by others. In studying the relationship of personality and leadership, this appears to be an important missing component.

An aspect of personality in many definitions is the matter of individual uniqueness, demonstrated by Bernard (1974) in his description of personality. "Personality is the unique and rather predictable patterns of orientation and response to biological, psychological, and other social factors that influence one's impact on others on others" (p. 6).

Personality and Leadership

A follower's definition and view of personality may contribute to a positive or negative perception of a leader. Birnbaum (1989) places personality and leadership into organizational perspective: "Leaders accumulate power through their offices and their own personalities to the extent that they produce the expected rewards and fairly distribute them. . ." (p. 23). Duke (1987) states "the effectiveness of school leaders depends on many factors--personality, reputation, imagination, courage, credibility, luck and dozens more" (p. 259). Personal recognition is also a factor for effective school leaders (Duke, 1987). Other personal qualities include resilience in the

ERIC

face of personal misfortunes and knowing that effort rules over inspiration.

Personality also relates to other qualities. Bennis (1989) contends that a leader values other people and has the personality to relate, develop, involve, and empower those working with and for the leader. Individual differences in personality and values play a role in whether transformational or transactional leadership emerges in a given situation. Sergiovanni (1990) explains:

> In 1978 James MacGregor Burns proposed a theory of leadership that has shaped new understanding of leadership practice. According to Burns, leadership is exercised when persons with certain motives and purposes mobilize resources so as to arouse and satisfy the motives of followers. He identified two broad kinds of leadership, transactional and transformative. Transactional leadership focuses on basic and largely extrinsic motives and needs; transformative, on higher order, intrinsic, and, ultimately, moral motives and needs. This later point is important to understanding Burns' theory. Transformative leadership is first concerned with higher order psychological needs for esteem, autonomy, and self-actualization and, then, with moral questions of goodness, righteousness, duty, and obligation (p. 23).

Personality contributes to meeting the needs of followers to a degree and in what ways that are still to be determined.

Psycho-historians have studied the lives of charismatic leaders. Individual personality and its development accounted for leaders' unique natures and styles. For example, the early death of a father (Franklin Roosevelt, Joseph Stalin) or the strong mother (Douglas MacArthur, Dwight Eisenhower, Harry Truman, Winston Churchill, and Napoleon Bonaparte) shaped sons, as directed by their mothers strong need for achievement, power, or recognition that could only happen through their sons' successes (Bass, 1981).

Leaders have inner strength and talents that can carry them in times of troubles and adversity. Moses, Buddha, Jesus, and Mohammed withdrew during difficulties and returned with strengthened faith in self and one's teaching. Speculation is that leaders who are more frequently transformational are likely to be higher in social boldness, introspection, thoughtfulness, and general energy but not sociability, cooperativeness, and friendliness than leaders who are more frequently transactional. Further research on the personality should focus on how important personality is to transformational, in contrast to transactional, leadership (Burns, 1978).

ERIC

Bass (1981) considers the problem of leader personality traits and their acceptability to a group that is highly homogeneous. Further research in isolating the factors that facilitate retention of a position of leadership, once it has been attained, may yield more about personal characteristics considered basic for effective leadership. Bass (1981) lists several personal characteristics associated with emergence as a leader: rate of talking and interaction, interaction capacity, task ability, dominance, exclusive possession of information, initiation of spontaneity, provision of group freedom, and acceptance of group members.

In summary, personality is a factor in a leader's approach to followers (Bass, 1981). Leaders use personality to engender trust and develop support for personal and organizational goals. Leaders' personalities and personal traits have an impact on their success with a given group of followers. In the search for leaders, personality considerations matched with follower expectations may promise more success than looking for any particular personality type.

Traits. The concept of traits has fascinated scholars for centuries. Whether a trait exists or not has yet to be proven to everyone's satisfaction; nevertheless, much thought has been

devoted to the study of the concept of trait and of various traits.

Definition

What constitutes a trait? The concept is a useful one; however, controversy even exists about the definition. Allport (1966) enumerated eight criteria about a trait:

1. Has more than nominal existence.
2. Is more generalized than a habit.
3. Is dynamic, or at least determinative in behavior.
4. May be established empirically.
5. Is only relatively independent of other traits.
6. Is not synonymous with moral or social judgment.
7. May be viewed either in the light of the personality which contains it, or in the light of its distribution in the population at large.
8. Acts, and even habits, that are inconsistent with a trait are not proof of the nonexistence of the trait (Allport, 1966, p. 3).

Criticism of traits. Critics of the concept of trait have argued that statements about a trait, such as "So and so is a nail chewer because of a nervous habit," is little more than a redundant description rather than a reference to causality.

Reification of a trait is also posed as a danger, as illustrated by these statements:

ERIC

Lynn behaves compulsively.

Lynn has a compulsive personality.

Lynn has a trait of compulsiveness.

What began with a description of how Lynn acts (on occasions) was fallaciously ascribed with misplaced concreteness, according to critics, to an inanimate object--compulsiveness.

Another argument against the concept of trait is that a person's behavior varies so much from situation to situation that sorting out varied behaviors into distinct categories is nigh impossible. The idea that a person can be "a street angel" but "a closet devil" is aptly descriptive of how one person can be twain. To some critics, personality traits are of less importance for study than the interactions among people which should provide a basis for identifying consistency of behaviors. For some critics, a trait is the residual effect of previous stimulation and reflects a person's adaptation level at the present.

Although many critics have pointed out the influence of situational factors as age, educational level, and other demographics, Allport (1966) reminds us that a person's outlook on life is still a part of that person. Situational differences may evoke varying responses; nevertheless, the responses come from an individual's reservoir of behaviors.

ERIC

Adult Personality Development

Personality development in adulthood has been treated with benign neglect. Early Formation Theories are abundant and have followed the pattern set by Freud that "as the twig is bent, the tree is formed." Early Formation Theories suggest that once past adolescence, nothing happens to a person's personality (Wrightsman, 1988).

Recent research on midlife crises has stimulated the study of adult personality development. Erickson (1959) says personality development in adults proceeds by stages with crises ending a stage; the stages build on each other so that the outcome of resolved crises affects a leader's ability and means to attack successfully the conflicts of the next stages. Personality development extends beyond biological and family ones; intertwined are the nature of society and its institutions with the stages of development (Wrightsman, 1988).

Changes in personality can occur in adulthood. Lawrence and Jaffe (1977) believe that evidence is beginning to accumulate that systematic and measurable changes can occur in adults. Can effective leaders be successful in multiple contexts by conscious use of experience and personality adjustment? There is a need for further research and a theory that will emphasize the ego or executive functions of personality.

Identity

The strength of one's identity contributes to the development of visionary leadership according to behavioral theories. A leader, as with any person, has the need to be recognized by others as an individual having identifiable qualities. Few individuals are self-sufficient, so that followers in a community contribute to a leader's self-concept and well-being (Barnhouse, 1984).

Interpersonal competence requires a strong identity to interact successfully with situations and followers. Interpersonal competence is observable behavior identified with health, intelligence, empathy, autonomy, judgment, and creativity. Leaders with interpersonal competence meet and deal with change, make strategic plans while executing previous plans, and enable other humans to make decisions. Interpersonal competence is a leader's capacity to make people capable of joint performance through common goals, common values, discovering new values, and new means while responding to change (Foote & Cottrell, 1955).

Personality Testing

A number of instruments purport to measure personality. Because of the changeability inherent in personality, there is probably no clear, solid, or fool-proof means of measuring it.

Researchers have slowed in designing new instruments for the measurement of personality. Instruments considered to be the best measurement of personality have been available commercially for some time.

Among test instruments available for personality are: (a) the Minnesota Multiphasic Personality Inventory (MMPI); (b) the California Psychological Inventory (CPI); (c) the Eysenek Personality Inventory (EPI); (d) the Personality Research Form (PRF); (e) the Edwards Personal Preference (EPP); (f) the Myers-Briggs Type Indicator (MBTI); and (g) the Human Resource Development Report (HRDR).

Minnesota Multiphasic Personality Inventory. This instrument, for ages 16 and over, is primarily for the measurement of personality. Developed by Hathaway and McKinley in 1940, the Minnesota Multiphasic Personality Inventory has been frequently updated. Reviewers indicated that "The MMPI, as the most extensively researched instrument in personality assessment, has a special responsibility to meet challenges confronting assessment as a whole" (Alker, 1984, p. 616), and "The Minnesota Multiphasic Personality Inventory (MMPI) is a standardized questionnaire that elicits a wide-range of self-descriptions scored to give a quantitative measurement of an individual's level of emotional

BEST COPY AVAILABLE

adjustment and attitude toward test taking" (Groth-Marnat, 1990, p. 179).

The MMPI has moderate levels of stability and internal consistency. Scales vary from .71 to .84, with scale intercorrelations high because of item overlap, which may lead to problems with scale interpretation. Studies support the construct validity of the MMPI (Groth-Marnat, 1990).

One caution in using the MMPI is the realization that in its original form it was designed for a two-pronged population, "normals" and psychiatric patients. "If as assumed, personality variables play a significant part in job satisfaction and performance, tests standardized on a normal population measuring variations along normal personality dimensions are far more appropriate than the pathology-oriented MMPI" (King, 1984, p. 936). Use of the MMPI has been widespread and it is well known by professionals in the field of personality.

<u>California Psychological Inventory</u>. Developed by Harrison Gough in 1957 and revised in 1987, the California Psychological Inventory has practical usefulness as its goal and strives to be simpler and easier to understand than the MMPI. In a survey of all clinical psychology programs in the United States and Canada with full American Psychological Association accreditation, 49% of the respondents included the CPI an

instrument with which Ph.D. candidates should be familiar (Piotrowski & Keller, 1984). These results placed it second only to the Minnesota Multiphasic Personality Inventory. The CPI is intended for use with "normal" individuals ages 13 and over, and the focus of the instrument is to measure elements of personal behavior that are to be found in all cultures and societies, and that hold a direct and important relationship to all forms of social interaction (Gough, 1968).

The purpose of the 18 scales is to predict what people will say and how they will behave in certain defined situations and to identify people who will be described and talked about in certain ways by observers and people who know them well. The instrument does not, however, claim to measure personality traits (Gough, 1968).

Reliability of the instrument is reported in the .20 to .50 range, giving it moderate levels of temporal stability and internal consistency. These coefficients do not appear on the surface to be substantial but, "Such relationships are typical in personality research, and extremely high correlations are unlikely to be found, since the scales are developed to assess rather broad behavioral tendencies" (Baucom, 1985, p. 251). In general the validity is good. The CPI measures such constructs as dominance,

self-control, and achievement, while the constructs of the MMPI are for use by more clinical populations and personalities.

The CPI is easy to understand and is good for measuring the constructs it is intended to measure (i.e., self-control, dominance, achievement). The CPI has been used extensively in areas of career development and personnel selection (Baucom, 1985).

Although the MMPI and the CPI appear to be similar, there are some significant differences between them (Groth-Marnat, 1990):

> Despite these similarities, it is essential for any clinician using the CPI to also appreciate the significant conceptual and psychometric differences between the two tests. The general intent of the MMPI is to assesses a person's intrapsychic processes and emotional distress as these relate to specific psychodiagnostic categories. Each of these categories has a group of internal dynamics surrounding it--such as depression, which also includes apathy, lowered capacity for pleasure, and feelings of hopelessness and helplessness. The primary task of the MMPI is to identify either the presence or absence of these internal dynamics and to place the examinee in either a normal or one or more psychopathological categories. In contrast, the CPI focuses more on a normal population and is highly interpersonal in

nature. In fact, there is a marked absence of symptom-oriented questions. Thus, the CPI is concerned with the presence or absence of specific interpersonal skills. In addition, the CPI avoids complex diagnostic nomenclature and instead emphasizes practical descriptions that are commonly used in most cultures (p. 238).

The CPI is a very useful instrument in the assessment of interpersonal characteristics of relatively normal persons. The CPI also measures variables of interest to a range of groups and individuals. The 1987 revision, which is closely related to the previous version, has not had extensive validity studies except in clinical settings. The absence of extensive validity studies must be viewed as a limitation of the instrument until further study is completed (Groth-Marnat, 1990).

Eysenck Personality Inventory. The Eysenck Personality Inventory is one of the best known self-report instruments on personality. This instrument measures two personality dimensions, those of extraversion (E) and neuroticisim (N), which consistently have appeared in the literature as being major traits. These two traits, extraversion and neuroticism, have been central to the theories of Eysenck, who developed this instrument in 1963 from the Maudsley Personality Inventory. The test-retest coefficients

ERIC

are .81 and .85 for individual scales and in parallel form .80 for neuroticism and .75 for extraversion (Tellegen, 1978).

This inventory is convenient and easy to administer. In the Eighth Mental Measurements Yearbook, Tellegen indicates that, "By themselves the N and E scales do not tell us enough for most purposes; but a person's standing, on these two dimensions provides a point of departure or baseline against which scores on other, more narrowly focused but related, self-report scales may sometime be evaluated more meaningfully" (1978, p. 803). Because of the broad range of scales, some overgeneralization may occur in scale interpretation. Although the EPI remains in print, an updated version known as the Eysenck Personality Questionnaire is also available.

Personality Research Form. The purpose of this instrument by Douglas Jackson is to assess personality traits and to measure the range of normal social and normal interpersonal behaviors. The theory for this instrument was derived from the work of Henry Murray and his list of needs. Murray emphasized the biological roots as well as the social and environmental determinants of behavior. He was also consistently aware of how individuals interact with their environment. This interaction includes how people are affected by outside forces and how their

ERIC

unique set of needs, attitudes, and values influences their reaction to the world around them.

There are extensive norms provided, but the samples are not well-defined, making comparison with other instruments difficult. The high content validity and homogenity make possible, ". . . the measurement of personality traits with levels of precision and validity formerly associated only with intellectual abilities and scholastic achievement" (Kelly, 1972, p. 298).

The Personality Research Form measures 15 constructs on the standard form and 22 on the long form. Dimensions measured are similar to those in other standardized inventories, i.e., achievement, dominance, and social recognition. Reliability scores are reported to range between .69 to .90 on test-retest, .54 to .86 on K-R coefficients, and .48 to .90 on odd-even (Anastasi, 1972, p. 298).

The PRF is a well-constructed test instrument, but the empirical validity has yet to be demonstrated. Many of the perceived shortcomings are characteristic of many instruments in personality assessment (Hogan, 1978).

Edwards Personal Preference Schedule. The Edwards Personal Preference Schedule was designed to measure 15 needs selected from Murray's list of manifest needs. The 15 scores measure 53 personality variables, including achievement, dominance,

affiliation, change and endurance. The editors of the manual suggest that the primary purposes for the development of the instrument were for research and counseling and to provide a measure of variables within the normal personality (McKee, 1972).

Norms were based on male and female high school and college students, and its intended population is college students and adults. While the validity scales are satisfactory, validity was not emphasized in scale construction, and there is little evidence to support it (McKee, 1972).

Generally reviewers have found the EPPS less useful for measuring personality variables than instruments like the CPI. The absence of supportive validity studies and low reliabilities also contribute to this perception.

In an updated version of the EPPS, the Edwards Personality Inventory (EPI) attempts to solve the problems encountered with the previous instrument. While the Edwards Personality Inventory has many innovative features, such as a multiscale approach, more data are needed concerning its stability over time.

In his review Norman (1972) suggests:

> . . . the EPI is an instrument worthy of serious consideration by those interested in the assessment of a broad range of personality characteristics in 'normal' adolescents and adults. I believe its usefulness will continue to grow as time passes and research evidence accumulates on the

ERIC

relationships of inventory scores and score profiles to relevant behavioral criteria of diverse sorts (p. 154).

Myers-Briggs Type Indicator. The Myers-Briggs Type Indicator is based on the theories of C. G. Jung. By self-report, the identification of reactions to preferences regarding perception and judgment is made possible. Ideally, this information can be used by a respondent to recognize and understand personality type of self. The MBTI is based on four bipolar constructs: (a) Extraversion vs. Introversion; (b) Sensing vs. Intuition; (c) Thinking vs. Feeling; and (4) Judgment vs. Perception.

When these scales are interpreted as continuous, their validity is good. While the format is forced-choice, each question deals with only one polarity. Correlations between the corresponding dimensions are high and statistically significant. Self-ratings of type and assignment are much closer than would be expected by chance (Wiggins, 1989).

While this instrument has been widely used by business, industry, and education, the greatest weakness of the MBTI is the inability to sustain the results over a period of time (reliability). On test-retest reliability for periods ranging from 5 weeks to 6 years, the statistics reported seldom exceed .50. The instrument is of value to normal populations for

identification of personality types at a given time (Wiggins, 1989).

Human Resource Development Report. The HRDR is based on the 16 Personality Factor Questionnaire, an adult personality inventory, and is designed to help persons have a better understanding of themselves in a managerial role. Major parts of the HRDR include: Leadership, Interaction with Others, Decision-Making Ability, Initiative, Personal Adjustment, and Distortion; the latter is a special scale constructed to detect faking responses. A library of statements is available to provide an analysis for patterns of scores. Although an ideal profile is not specified, there are norms, and the statements give a clear description of behaviors and activities in the areas. Nevertheless, how to interpret the results is left up to the examinee. A test-retest reliability for two weeks is given as .81. The five major parts were determined to be recurring themes from a review of literature, but evidence on how scores relate to performance is unavailable. Whether personality instruments can predict specific behavior is open to question.

There is a range of personality measurement instruments available in addition to the ones reviewed. Each should be evaluated by the needs of the administrator and population to be tested. Additional test instruments warranting investigation by

those interested in personality measurement and identification include: (a) Jackson Personality Inventory, (b) Rorschach, (c) Millon Clinical Multiaxial Inventory, (d) Thematic Apperception Test, and (e) Projective Drawing Techniques.

Leadership

The Study of Leadership

What is leadership? Is it measurable? Can it be isolated for study? Can it be developed? Is it a trait? Is it a function of situational factors? Is it a birthright? Is it the result of a host of demographic factors?

Probably more is written and less known about leadership than any other topic in the Behavioral Sciences (Bennis, 1989). What is known can be valuable to the researcher and the practitioner.

Definition. To some, leadership is considered to be personalistic, composed of a set of characteristics or traits. To others, leadership is a process of group/member interactions. Another view is that the role behavior of persons in status positions best defines leadership. Leadership can be simply thought of as the process of moving a group to the accomplishment of some goal without force. Leadership is also used to refer to persons in roles where leadership (the process of moving a group . . .) is their responsibility.

> Everyday usage of the phrase "good" or "effective leadership" has even more meanings, but most stress one key point. "Good" leadership moves people in a direction that is genuinely in their real long-term best interests. It does not march people off a cliff. It does not waste their scarce resources. It does not build up the dark side of their human nature. In this sense, one could say Adolf Hitler displayed strong leadership at times, but obviously not effective leadership (Kotter, 1988, pp. 16-17).

How leadership is defined or viewed has great impact upon how "good" or "effective" one's view of another's leadership is rated. When Harry Truman was a Senator from Missouri, his leadership at the national level was limited. After he became President, many national observers, even some of Truman's detractors, acknowledged his successes in the Oval Office. Truman, on the other hand, thought of his successor, Dwight Eisenhower, as a sit-still president, surpassed as this kind in our century only by Coolidge (Truman, 1989).

Leadership is a difficult concept and consumes the attention of scholar and practitioner alike. In the early part of this century, the belief was that leaders were born, not made. Research since that time has progressed under major six categories. The categories are not specifically sequential or

fixed but rather fluid; they are a way to think about the many theories and models that have emerged. The categories include trait theories, power and influence theories, behavioral theories, contingency theories, cultural and symbolic theories, and cognitive theories.

Trait theories. Trait theories "attempt to identify specific personal characteristics that contribute to a person's ability to assume and successfully function in positions of leadership" (Bensimon et al., 1989, p. 8). Traits may include physical characteristics, personality, social background, and ability. A combination of traits does not guarantee effectiveness even though they appear to be characteristic of successful leaders (Bass, 1981). Few, if any, traits are essential for successful leadership (Bass, 1981; Gibb, 1968).

Power and influence theories. Social power theories emphasize one-way influence while social exchange theories focus on reciprocal relationships between leaders and followers.

> Leaders accumulate power through their positions and their personalities, but their authority is constrained by followers' expectations. In essence, the group agrees to collectively reduce its own autonomy and to accept the authority of the leader in exchange for the rewards and benefits (social approval, financial benefits, competitive

BEST COPY AVAILABLE

advantage) the leader can bring them (Bensimon et al., 1989, p. 15).

Leadership is related to followers' expectations. To be successful, leaders must either fulfill these expectations or change them. This is the distinction between transactional and transformational leadership (Bass, 1985; Bennis & Nanus, 1985; Burns, 1978; Bensimon et al., 1989). Transactional leadership is a relationship between leaders and followers that is like a system of barter. The leader and follower exchange human needs and interests that can be economic, political, or psychological in nature. Change continues as long as the bargain is agreeable and success depends on values like honesty, firmness, and honoring commitments (Bensimon et al., 1989).

Transformational leadership takes the next step and is a "building" approach (Sergiovanni, 1990) that helps followers to grow to greater levels of morality, motivation, potential, and expectations. Sergiovanni suggests that transformational leadership as bonding is giving purpose, meaning and significance to followers, and the effects are performance and commitment sustained beyond expectations in quantity and quality. Bonding stresses cultural and moral leadership that lifts organizational goals and purposes to the level of a shared covenant and bonds together leader and followers in a moral commitment. Personality

of a leader affects collegiality while providing the means to move people from subordinates to followers. "Subordinates respond to authority; followers respond to ideas" (Sergiovanni, 1990, p. 27). The next step, according to Sergiovanni is leadership as banking. Improvement in the organization becomes routine when followers and leaders share a covenant. Here the leader is "ministering" to the needs of the school by being of service to others so they are better able to perform. The leader guards the values of the organization. The leader practices servant leadership in which the leader "builds up the leadership of others and who strives to become a leader of leaders" and "bureaucratic authority and psychological authority are transcended by moral authority" (Sergiovanni, 1990, p. 27). Further, the leader must have a personality and identity capable of listening, intervening appropriately, serving, trusting and being relational.

Sergiovanni (1990), referring to James MacGregor Burns' (1978) theories on transactional and transformative leadership, suggests a leader must know when to use each of the stages of bartering, building, bonding, and banking. Further, a leader must have the personality and the identity to listen to the process, intervene appropriately, be a servant, trust, and be relational.

Behavioral theories. A behavioral approach accumulates data about leaders through observations, questionnaires or

ERIC

self-reporting devices. This theory considers what leaders actually do (Mintzberg, 1973). Early studies analyzed the effects on a group's performance compared with a particular style of leadership. Researchers looked at authoritarian, democratic, and permissive styles while studying the methods that leaders used. The primary research was conducted as part of the Ohio State Leadership Studies. An instrument developed to assess a person's leadership style, the Managerial Grid, included a two-dimensional array with two scaled axes (Blake & Mouton, 1964). The instrument was designed to identify "initiating structure" or task orientation and "consideration" or relationship orientation (Stogdill & Coons, 1957). Critics of the instrument suggest that the best leadership style is to be high on both orientations without looking at the task, the environment, or the level of the followers (Bensimon et al., 1989).

Research on behaviors resulted in identification of managerial roles. The roles fall under three groups: interpersonal behavior, information-processing behavior, and decision-making behavior (Mintzberg, 1973).

These behavior theories have run into difficulty because an effective leader engages in those behaviors that are applicable for a specific situation and changes as the followers or context changes. Linking the performance of a group to the behaviors of a

leader has been problematic in determination of the cause of the performance. Predicting outcomes in advance of leader behavior with group effectiveness has been difficult (Bensimon et al., 1989).

Contingency theories. An effective leader changes approaches and behaviors according to the situation, the type of group, its level of ability, the type of task, and the environment, according to contingency theories. Behavioral theories and the contingency theories "overlap" with the former focusing on internal issues and the later on external environment. The theories suggest that no single leadership approach is best but not all are equal in effectiveness (Bensimon et al., 1989).

Fiedler's (1967; 1971) contingency model is one of the more well known approaches and uses the task, or relations-oriented, approach. Fiedler added a leader's intelligence, competence, and experience (Fiedler & Garcia, 1987) and calls it the Cognitive Resource Theory. Another contingency model is the Situational Leadership Theory by Hersey and Blanchard (1977).

The theory relates behavior of leaders to the maturity of followers. House (1971) proposed the Path-Goal Theory that suggests a leader clears the "path" to goal attainment, thereby increasing follower satisfaction.

ERIC

Vroom and Yetton (1973) assert that the level of participation by followers in decisions may determine the effectiveness of a leader to reach the leader's desired outcomes. Yukl (1981), in the Multiple Linkage Model of Leader Effectiveness, suggested that environmental factors and group dynamics may deter, in the short term, the achievement of the leader's or the group's desired outcomes.

Organizational characteristics may be substitutes or neutralizers of leadership (Howell, Dorfman, & Kerr, 1986; Kerr, & Jermier, 1978). For example, the culture of a group may be built on tradition which prevents any change beyond the boundaries set by tradition and, in effect, the tradition is the leader. Kerr and Jermier (1978) studied "the nature of situations in which neither task or (sic) consideration leadership may have any effect on subordinates' satisfaction, motivation or performance" (Bensimon et al., p. 19).

Contingency theories have had mixed reviews and provide a primarily environmental view of leadership. These theories assume that leaders and organizations are rational and linear. The authors of cultural and symbolic theories look at leadership from the perception of the leader.

Cultural and symbolic theories. These theories "assume that organizational structures and processes are invented, not

discovered. . . These theories propose that leadership functions within complex social systems whose participants attempt to find meaningful patterns in the behaviors of others so that they can develop common understandings about the nature of reality" (Bensimon et al., 1989, p. 21). Cultural and symbolic theories are used to study a leader's view of organizational data and how a leader processes the data.

The interpretation of culture and its development is the focus of the leader (Deal & Kennedy, 1982). By developing new symbols, stories, myths, and legends, a leader establishes and reinforces consistent values that result in an increased commitment to an organization, motivates participants, and leads to organizational excellence.

Leaders may be able to change attitudes of participants in organizations but may not have much effect over the outcomes of organizational behavior (Pfeffer, 1981). Birnbaum (1989) studied presidential leadership of colleges and universities over ten years and contends that little relation exits between changes in the presidency and institutional functioning. Leaders attending to organizational culture have little effect on organizational goals.

Cognitive theories. Cognitive theories of leadership emerge out of research on organizations. Cognitive theories consider

that leaders look at organizations to find meanings in unusual occurrences (Bensimon et al., 1989). Leaders are to cause change. Leadership is associated with sets of myths that help followers to believe in leaders' effectiveness in causing outcomes (Fiedler & Garcia, 1987). Cognitive processes that give attention to leaders' personal decisions permit leaders to claim credit for positive outcomes because of their own ability and effort. In like manner, when leaders' decisions go awry, they can shift the blame for failures to external causes such as luck or the difficulty of the task (Bensimon et al., 1989). As long as leaders' interpretations of events and decisions are accepted by followers, "Successful leaders . . . are those who can separate themselves from organizational failures and associate themselves with organizational successes" (Pfeffer, 1977, p. 110). Therefore, "cognitive theories regard leaders as an invention of followers. What matters is perception: If leaders are seen doing the desired leaderlike things they will be regarded as effective leaders" (Bensimon et al., 1989, p. 26).

Trait theories are the oldest while cognitive theories are some of the newer approaches to a study of leadership. Personality finds its most visible influence in trait theories and power and influence theories. Traits, distinguishing qualities or characteristics, continue to show up as a factor in

ERIC

leadership. For example, behavioral theories look to traits as a factor in formulation of "styles."

Scholars' work on leadership has also been grouped under three main approaches (Lipham & Hoeh, 1974): psychological, sociological, and behavioral.

Psychological approach. In this approach to the study of leadership, what a person "is," or what comprises basic personality structure, is as important as what a person does or is expected to do. The "Great Man" theory contends that what makes a leader can be learned from studying the lives of great men--who were born so and not made. The lives of political figures, military commanders, and great industrial barons are three fruitful sources for study.

In education, the literature on effective schools, with emphasis upon the importance of the principal, calls to mind the prominence of individuals. In many schools, colleges, and universities, walls are adorned with pictures of the great who have held the principalship, deanship, or other administrative position. For some, the pictures serve as a motivation to become successful and attain a status position like those pictured.

Another avenue of study is the identification and measurement of leadership traits, often couched as personality traits. Self-report inventories and other measures are used to examine

desirable traits. In some cases, traits are seemingly mutually exclusive, as a leader must be "flexible but firm." Many measures have been used to study personality characteristics and have shown that individuals in leadership positions scored higher than their subordinates in many instances. From several studies, for example, supervisors have demonstrated greater accuracy in perception, more initiative, greater dominance, higher achievement and mobility drive, and greater tendency to identify with superordinates (Lipham & Hoeh, 1974). While the study of personality traits has not produced a universal set of traits, some patterns of traits seem to distinguish leaders from nonleaders.

Sociological approach. In the absence of conclusive results from the psychological approach, the sociological approach was devised to study social relationships and roles, particularly with groups, for a determination of leadership. The work of Hemphill and others at Ohio State University focused upon the differences among groups and the effect of leaders upon them. Group members and their leaders have been studied through observations, interviews, simulation exercises, and decision-making activities. While situational factors contributed to further understanding of leaders and their behaviors, a realization that leadership was

ERIC

somehow broader than situational factors led to yet another approach.

Behavioral approach. This approach examines the complex interrelationships of personal traits, situational factors, and behaviors. While examining the observed behaviors of leaders-in-situations, followers of this approach differentiate between leader behavior and leadership. Focus is directed upon observed behavior rather than upon preconceived capacities or abilities. Further, leader behavior is not presumed to be either innate or situationally directed. Care is taken to distinguish between a description of leader behavior and an evaluation of it. One result of the behavioral approach is that leadership is viewed as a highly complex set of interrelationships.

Leaders and Leadership

Leadership, as defined by Hersey and Blanchard (1977), is a process of influencing the activities of an individual or a group in efforts toward goal achievement in a given situation. Certainly an educational leader hopes to influence teachers and students toward quality education. How well a leader achieves this hope effects learning outcomes. What, then, are the qualities of an effective leader?

Vision. Bennis (1989) calls vision a dream that leaders manage. Bass (1985) sees vision as an image of a desired future

organizational state. An alternative image of visionary leadership might be of a drama. "Idea and emotion, actor and audience, are momentarily united in a rich encounter which occurs on many symbolic levels" (Westley & Mintzberg, 1989, p. 18). Erikson (1978) suggests leaders who inspire such a world view are called "eminent" and even considered to be "immortal." Leaders who seek human brotherhood in self-denial have another world-view vision. They are leaders who know why they were put on this earth.

The difference between more and less effective leaders may boil down to two factors: More effective leaders (a) possess--and are able to articulate--a vision of effective schooling and (b) allocate their time in ways that increase the likelihood of realizing that vision (Duke, 1987). Effective leaders work to gain the trust of their constituents, communicate their vision lucidly, and involve everyone in the processes of change (Bennis, 1989).

Leaders distinguish themselves from the general run of managers in at least six respects according to Gardner (1990):

1. They think long term, beyond the day's crises.

2. They grasp the connections between their unit and the larger organization, external conditions, and global trends.

ERIC

3. They reach and influence constituents beyond their jurisdiction, beyond boundaries.

4. They put heavy emphasis on the intangibles of vision, values, and motivation and understand intuitively the nonrational and unconscious elements in leader-constituent interaction.

5. They have the political skill to cope with the conflicting requirements of multiple constituencies.

6. They think in terms of renewal. The leader or leader/manager seeks the revisions of process and structure required by every-changing reality (p. 4).

Effective leaders with vision develop a commitment to their organizations. Sergiovanni (1990) suggests school improvement occurs best under transformational leadership with a style he calls "bonding." When a leader satisfies a follower's desires for purpose and meaning through elevation of organizational goals and purposes to the level of a shared covenant, then the follower's performance is beyond expectations and changes occur regardless of external conditions. A bonding occurs between leader and followers and translates into a moral commitment. The switch from bureaucratic leadership to moral imperatives, based on shared values, translates into servant leadership on the part of both leader and follower and the outcome is commitment and quality performance beyond the call of duty (Sergiovanni, 1990).

<u>Qualities of leaders</u>. People have thought that the heroic leader was born with the qualities for leadership. Alexander the Great and Winston Churchill are such examples. But many scholars agree with Bennis (1989a) that each person contains the capacity for leadership by becoming an integrated person. Leaders know who they are, what their strengths and weaknesses are, how to fully deploy their strengths and to compensate for their weaknesses. They know what they want, why they want it, and how to communicate what they want to others so that they gain cooperation and support.

Bennis' qualitative research subjects (90 leaders in their field) reject the notion that leaders are born but rather that leaders are made. He found that his subjects continue to grow throughout life and, therefore, taking care of their own learning is part of being an integrated person and a leader.

Many qualities are ascribed to leadership: Vision, independence from the context, character, drive, technical competence, people skills, conceptual skills (meaning, imagination, and creativity), judgment and taste (Bennis, 1989). Bennis adds these ingredients: A guiding vision stemming from desire to change the culture, passion for vocation, a profession, a course of action, integrity, trust, curiosity, and daring.

Scholars have yet to come to unanimity as to the qualities of leadership with one exception. Research on characteristics of effective leaders, in the private sector, finds one universal characteristic: vision (Bolman & Deal, 1990).

Relational leadership. An effective leader has rich relationships with followers, peers and superordinates. Leaders cannot "like" everyone, therefore leaders build relationships that can deal with differences. The first step is to separate relationship (process) issues from substantive ones. "The strategy is to be unconditionally constructive" (Fisher & Brown, 1988, p. 37). Fisher and Brown consider the following as the six elements of a working relationship with others:

1. Rationality: Even if they are acting emotionally, balance emotions with reason;

2. Understanding: Even if they misunderstand us, try to understand them.

3. Communication: Even if they are not listening, consult them before deciding on matters that affect them.

4. Reliability: Even if they are trying to deceive us, neither trust them nor deceive them;

5. Noncoercive modes of influence: Even if they are trying to coerce us, neither yield to that coercion nor try to coerce them;

ERIC

6. Acceptance. Even if they reject us and our concerns as unworthy of their consideration, accept them as worthy of our consideration, care about them, and be open to learning from them. (p. 38).

In a two-person relationship, a leader wants to be able to take steps that will improve working together and advance the substantive interests of both. By following the six guidelines, Fisher and Brown (1988) contend that putting these rules together and applying them appropriately and consistently produce a coherent and effective relationship. An effective relationship is "congruence," which occurs when performance and commitment are sustained beyond external conditions and limitations (Fisher & Brown, 1988).

Training

The lack of agreement on characteristics of effective leaders and theories of leadership makes training difficult. Schroeder (1986) reviewed training and found that the most common approach to leadership training and interpersonal skills has been textbooks and simulations. Role playing and assessment center simulations are effective. Both methods require a large number of instructors, including role players, evaluators and administrators (Schroeder, 1986).

ERIC

The U.S. Army Research Lab produced highly interactive videodisc training scenarios covering 20 leadership problems. The results indicated significantly greater learning of leadership principles. Role-playing was slightly higher than videodisc and both-role playing and videodiscs were much higher than a programmed text. A combination of videodisc and role playing were found optimal for leadership training (Schroeder, 1986).

A two-year study by Smith and Monahan (1990-91) looked outside of education and to the military because of its focus on leadership as an academic field. Schools of education and business have recently emphasized management. They found that the military theory is an interactionist approach. This approach is based upon the assumption that both the individual and the situation interact to produce effective leadership (Smith & Monahan, 1990-91).

Smith & Monahan (1990-91), in a two year study, looked for training models outside of education. The military approach proved interesting because of its focus on leadership as an academic field. They found that military theory is an interactionist approach which is based upon the assumption that both an individual and a situation interact to produce effective leadership (Smith & Monahan 1990-91).

ERIC

Smith and Monahan (1990-91) applied the military interactionist principles to education administration:

1. "Place the responsibility for organizational goal accomplishment on the leader to clearly differentiate between leadership and management" (p. 163-176). Often school boards and university regents micro-manage and inhibit a dynamic school administrator's leadership.

2. Measure outcomes, not by graduation rates, but by the "desired societal outcome and measure the education institution's success in meeting these goals" (p. 163-176). The military measures the leader's success in achieving the assignment or mission.

3. Formal training of university officials and school leaders could take place outside of the academy in a war college as part of their academic training.

4. "Hold persons in leadership positions to high ethical standards" (p. 163-176). Give them the authority and the responsibility to accomplish the goal. Relieve them if they are not effective in that context. Development of army leadership is explicit and systematic while effective leadership in education is mostly accidental. The interactionists believe that the trained manager only becomes a leader under "simulated" conditions in training and "live" in the situation (Smith and Monahan, 1990-91).

ERIC

Smith and Peterson (1988) report "the most recent view is that intelligence and possibly other personal qualities do play a quite substantial role in the emergence of leaders in unstructured settings" (p. 5). The majority of recent studies in "unstructured settings" are through assessment centers, which set up simulations of settings potential leaders are likely to encounter in their work.

Let us look at training of leaders from an event management model. The model proposes to bring together past research focusing on the social context of leadership actions that gives them their meaning and consequently their effect. "Leadership which contributes to effective event management can be defined as actions by a person which handle organizational problems as expressed in the events faced by others" (Smith & Peterson, 1988, p. 80). This implies that leadership is not just done by those appointed to it and that leadership influence not only moves from the more powerful to the less powerful in an organization.

Training will only be successful when it has a clear purpose that is shared by those being trained. The method suggested by Smith and Peterson (1988) is based on training for choice. "Leaders therefore often need to learn how to choose which sources of information or demands upon them to attend to, and how to choose which specific behaviors would stand the best chance of

implementing their intentions" (Smith & Peterson, 1988, p. 165). Awareness of choices in day-to-day operations is basic to good leadership. Further, "Exercises should focus on diagnosing one's goals and strategies, and then move on to the more specific level of how one might implement choices within the existing demands and constraints" (Smith & Peterson, 1988, p. 165). Training should be in the context of the specific culture or organization and emphasize that leaders and their contexts constantly redefine one another, rather than coexist (Smith & Peterson, 1988).

Leadership in business. A prime source on administrative and organization theory and practice is the world of business. What constitutes leadership in the business world? In business, a senior manager should meet the following requirements (Kotter, 1989):

I. Industry and Organizational Knowledge

Broad knowledge of industry (market, competition, products, technologies)

Broad knowledge of the company (the key players and what makes them tick, the culture, the history, the systems)

II. Relationships in the Firm and Industry

Broad set of solid relationships in the firm and in the industry

ERIC

III. Reputation and Track Record

Excellent reputation and a strong track record in a broad set of activities

IV. Abilities and Skills

Keen mind (moderately strong analytical ability, good judgment, capacity to think strategically and multidimensionally)

Strong interpersonal skills (ability to develop good working relationships quickly, empathy, ability to sell, sensitivity to people and human nature)

V. Personal Values

High integrity (broadly values all peoples and groups)

VI. Motivation

High energy level

Strong drive to lead (power and achievement needs backed by self-confidence) (Kotter, 1989, p. 30)

How does this list of requirements compare with those of school administrators at comparable levels? with those at lower levels? Some words, as "products," might not apply while "integrity" and "strong interpersonal skills" are no less desirable in school administrators.

ERIC

Leadership Testing

Leadership is defined by Fiedler (1987) as "the degree of success with which a group performs the primary task assigned" (p. 4). While there is no shortage of definitions as to what constitutes leadership, there is a shortage of instruments which will accurately present factual evidence related to leadership identification and selection. The identification and selection of leaders is at best risky. Although there has been substantial success in identifying leadership behaviors and traits, the prediction of success is viewed as somewhat more difficult. Interest continues, however, since 1980 there have been many studies focusing on the study of leadership and its measurement.

Many instruments have been developed to measure leadership. The instruments vary in assumptions about the nature of leadership and their subsequent approach to sampling it. Some of the better known instruments are described as an indication of the variety of available instruments and the ways that information on leadership is acquired.

The following instruments are among those available to persons desiring to obtain some indication and measurement of leadership: (a) Leader Behavior Description Questionnaire, Form 12; (b) Educational Administrator Effectiveness Profile; (c) Least Preferred Coworker; (d) Educational Administrator Effectiveness

Profile; (e) Situational Leadership; (f) Behavioral Assessment; (g) Administrator Perceiver Interview; (h) Leadership Skills Inventory; (i) Managerial Skills Inventory; (j) Managerial Philosophies Scale; (k) Purdue Rating Scale for Administrators and Executives; (l) Self-scoring Survey of Educational Leadership; (m) Styles of Leadership Survey; and (n) Styles of Management Survey.

Leader Behavior Description Questionnaire Form 12.

Developed by Halpin and Winer, Form 12 of the Leadership Behavior Description Questionnaire grew out of the original Ohio State research on the Leadership Behavior Description Questionnaire and the Supervisory Behavior Description Questionnaire. These instruments have been in use since the late 1950s and have proven to be very effective in identifying the behaviors of leaders. However, the results are not intended to be predictive in nature (Dipboye, 1978).

LBDQ-12 differs from the earlier two instruments in the number of dimensions measured. Form 12 measures twelve dimensions, while in the Leadership Behavior Description Questionnaire and the Supervisory Behavior Description Questionnaire only two dimensions are measured, initiation of structure and consideration.

Initiating structure measures a leader's behavior in determining the relationship to the group and recognizes patterns

of organization, communications, and procedures. Consideration measures the behaviors of friendship, mutual trust, and the relationship between the leaders and the group (Lipham, Rankin, & Hoeh, 1985).

Initiation of structure and consideration are still the most commonly used scales and "have been found to correlate significantly with satisfaction and performance, although the direction of these correlations varies with situational differences" (Dipboye, 1978, p. 1750). Dipboye also reports test-retest reliability scores range from .52 to .72 for initiation of structure (task) and .71 to .79 for consideration (relationships). Internal consistency ranges from .38 to .91.

The LBDQ-12 appears to have concurrent validity because the scales "have been found to correlate with the external criteria of job satisfaction and performance and are capable of distinguishing between persons displaying behaviors corresponding to the dimensions" (Dipboye, 1978, p. 1751).

Scoring for the twelve dimensions is based on employee ratings of supervisors and are evaluated by 4-10 raters. The following dimensions have been shown to have the highest interrater correlation: (a) Demand Reconciliation, (b) Tolerance of Uncertainty, (c) Persuasiveness, (d) Role Assumption, (e) Predictive Accuracy, and (f) Superior Orientation.

Representation, Tolerance of Freedom, and Integration have the lowest interrater reliability (Dipboye, 1978).

Perhaps the best use of this instrument is for research or the evaluation of leadership training programs. Although a lack of definitive norms lessens its statistical appeal, the LBDQ-12 is still regarded by many as the cornerstone instrument for the identification of leader behaviors.

Educational Administrator Effectiveness Profile. The purpose of this instrument is to rate the effectiveness of an administrator through self-evaluation and the evaluation of five coworkers that are qualified to evaluate the administrator's work. Norms are based on a sample of 195 administrators and 462 other raters. Constructs evaluated in this instrument include those generally associated with different administrative tasks, i.e., setting goals and objectives, making decisions and solving problems, and building and maintaining relationships (Hughes, 1989).

Descriptive statistics indicated a limited ability of the instrument to discriminate among administrators who vary in effectiveness (Hughes, 1989). The inter-rater reliability is moderate for the 11 scales (.50) and the range for internal consistency is .71 to .86. Ninety-three percent of the items correlate higher with their own scale totals than with other scale

totals. "Additional evidence of the construct validity is provided by correlations between self-ratings and ratings by others for each of the 11 scales. Generally these correlations are low, with higher correlations for task-oriented skill domains and lower correlations for person-centered skill domains (p. 281).

The EAEP can be most effectively used for organizational development and for encouraging the professional growth of administrators. There appears to be little relationship between the EAEP scores and external criteria of effectiveness (Hughes, 1989).

Least Preferred Coworker. Although the Least Preferred Coworker is a social-psychological test, it has been included in the leadership section because it measures leadership styles through a simple personality test. This instrument, designed by Fiedler, was part of the research to test his contingency model. The 16 test items of descriptive personality adjectives are based on a bipolar eight-point scale, and the score on these items suggests whether the respondent is task-oriented or relationship-motivated. Responses are generated by asking the person being tested to recall characteristics of someone with whom the respondent least likes to work, not necessarily someone that is disliked. Questions are then answered within this frame of reference (Fiedler, 1987).

A low LPC score (63 or less) indicates a person who has a high need to accomplish the task; and a high LPC score (73 and above) indicates a person who is relationship-oriented and willing to accept the personalities of others. Based on 898 cases the mean score was 68.75 with a standard deviation of 21.8. These cases represented a wide variety of population samples. Split-half statistics for reliability ranged from .86 to .92. Factor analysis generally showed a major item factor describing personality and a small factor describing task-relevant behavior (Fiedler, 1987).

The meaning of LPC scores is still a matter of some debate, and the interpretation of what the LPC measures has changed over the years. As Hoy & Miskel (1978) point out:

> At first, it was seen as a measure of an emotional reaction to individuals with whom the leader found it difficult to work; then, it was thought to differentiate between individuals who had a task orientation as opposed to an interpersonal one, or to measure cognitive complexity; today it is taken as an indicator of the motivations of the leader (pp. 286-287).

Situational leadership. Situational leadership is a blanket term for six instruments designed for different aspects of leadership in given situations. These instruments include: (a) Interaction Influence Analysis; (b) Leader Effectiveness and

ERIC

Adaptability Description; (c) Maturity Style Match; (d) Maturity Scale; (e) Power Perception Profile; and the (f) Leadership Scale.

Of these six measurement instruments, the Leader Effectiveness and Adaptability Description (LEAD) is the oldest and best developed of the six. LEAD comes in two forms, the LEAD-self and LEAD-other. LEAD-self contains 12 leadership situations in which a respondent evaluates self and measures the leader behaviors of (a) style; (b) style range; and (c) style adaptability. The LEAD-other contains the same 12 leadership situations and is designed to provide a perspective other than one's own perception of leadership style.

Suggested uses of the Situational Leadership instruments are for managers, administrators, leaders, supervisors, and staff. Scale ratings are obtained by scores from observers, self-report, and others (Eberhardt, 1978). Theory for the instrument is based on Paul Hersey and Kenneth Blanchard's Situational Leadership Theory (Tri-Dimensional Leader Effectiveness Model). "Situational Leadership Theory is based on an interplay among (a) the amount of direction (task behavior) a leader gives, (b) the amount of socio-emotional support (relationship behavior) a leader provides, and (c) the 'maturity' level that followers exhibit on a specific task, function, or objective that the leader is attempting to

accomplish through the individual or group (follower[s])" (Eberhardt, 1978, p. 1384).

Developed from studies conducted at the Center for Leadership Studies, situational theory identifies four styles of leadership:

1. High task-low relationship (telling).
2. High task-high relationship (selling).
3. Low task-high relationship (participating).
4. Low task-low relationship (delegating).

Norms for these instruments were established by using 264 managers, ages 21 to 64, and the validity scores ranged from .11 to .52. There was a correlation of .67 between the adaptability scores of the managers and the independent ratings of their supervisors. Reliability was found to be moderately stable. Over a six-week period, 75% of the managers maintained their dominant style and 71% maintained their alternative style (Eberhardt, 1978).

While the situational leadership theory has many appealing aspects, Zedeck (1978) suggests that use of the instruments should be limited to research purposes and results interpreted in the context of the situation in which a leader will actually perform. Of particular concern with this set of instruments is the interpretation of the scores and the lack of specific recommendations as to their most productive use.

BEST COPY AVAILABLE

Behavioral assessment. Behavioral assessment is both a set of specific techniques as well as a way of thinking about behavior and how behaviors can be changed. It differs from other means of traditional assessment by attempting to understand situational determinants of behavior. Behavioral assessment is more concerned with cause and effect and may be a more direct and functional form of measurement. This subject is presented in greater detail elsewhere in this document.

Administrator Perceiver Interview. This instrument, developed by Selection Research, Inc., of Lincoln, Nebraska, is a structured interview composed of 70 questions on 14 themes, as Mission, Relator, and Delegator. The interview is individually administered by trained interviewers who are required to have 85% agreement in scores; internal consistency reliability is reported as .83. Information on content validity is not available. The purpose of the API is to provide information for employment decisions and professional development. Scores on the 14 themes are not to be thought of as "subtests" but only the total score is to be used to predict future behavior of interviewees. The API provides information on likely job-related characteristics of interviewees with emphasis on the building or positive administrator-teacher relationships and positive, open school climate. Data on the ability of the instrument to predict future

ERIC

behavior of interviewees or on its relationship to student or staff outcomes are not available. Because the 14 theme scores are highly correlated, the API does not measure different aspects of administration.

Leadership Skills Inventory. This inventory is self-administered and self-scored. It contains 125 items related to personal leadership skills for persons who aspire to improve their potential for leadership. The items cover 9 skill areas: Fundamentals of Leadership, Written Communication Skills, Speech Communication Skills, Values Clarification, Decision Making Skills, Group Dynamics Skills, Problem Solving Skills, Personal Development Skills, and Planning Skills. Hand scored, the results are to be compared with those of 452 persons from a comparison group. Scores below average should be thought of as areas for improvement. Internal consistency reliabilities range from .81 to .93; however, content validity was based upon the author's decisions about what constituted leadership skills from a review of the literature and is not supported by statistical analysis. For example, data on how the self-report on a skill area compares with another instrument would be useful so that a measure of concurrent validity could be obtained.

Managerial Philosophies Scale, 1975. This instrument asks respondents to agree/disagree with 36 statements to determine

disposition toward a Theory X or Y view of mankind. Persons who subscribe to Theory X believe that employees are basically lazy and motivated by close supervision, extrinisic rewards, and, if necessary, punishment. Advocates of Theory Y contend that employees are motivated by intrinsic and extrinsic rewards, enjoy work for itself, and can be self-directed. There is no evidence that the statements measure managerial approaches to subordinate motivation and organizational behavior that were propounded by McGregor (1960). No data on reliability are reported.

Purdue Rating Scale for Administrators and Executives, 1950. Subordinates rate their supervisors on 36 items for the purpose of self-improvement of their supervisors. Scores are to be interpreted normatively by an examination of the various ratings and the individual items. No test reliability is available. The items were developed from a review of the literature and from interviews with supervisors and subordinates about what constitutes effective administrative behavior. There is no information on validity, reliability, or usefulness for educators; therefore, its usefulness is questionable.

Self-Scoring Survey of Educational Leadership, 1979. Designed to assess leadership style, the instrument consists of 50 situations to which respondents are asked to indicate ideal and actual behavior of colleagues. Responses are interpreted into

four styles of leadership that are not named. Some averages for various groups are given, but none for educators. Technical information is lacking. The instrument is probably best used as an informal appraisal prior to training in leadership skills. Its usefulness to experienced school administrators is doubtful.

Styles of Leadership Survey, 1986. Designed for use in education, business and industry, this self-report instrument has 60 items to measure two dimensions--concern for people and concern for outcomes. These two dimensions are used in several instruments and were propounded as the Managerial Grid by Blake and Moulton (1964). [See the Styles of Management Survey.]

The instrument is self-administered and self-scored. From a comparison with nearly 3,000 persons, scores are placed on a two-dimensional grid that provide a profile of leadership style and strength of style. The differences between style scores show how heavily one relies upon the dominant leadership style and how much one uses the various styles. Feedback from others can be obtained by using The Leadership Appraisal Survey.

Information on both validity and reliability is limited in the manual. While the Managerial Grid serves as the theoretical base of the instrument, the linkage between the theory and the instrument is not provided.

Styles of Management Survey, 1986. Much like the Styles of Leadership Survey, this instrument purports to differentiate between concern for people and concern for outcomes. This instrument is also predicated upon the Managerial Grid. The instrument is self-administered and self-scored and consists of 60 items. The scores of 12,809 managers are the basis for comparing one's scores on leadership style and strength of style. Information from others can be obtained from The Management Appraisal Survey.

While information on reliability is limited to a reference to a coefficient of stability of .72, the authors assert that (a) the instrument discriminates between high, average, and low managers although the criteria are not identified, (b) the scores correlate with a personality inventory, and (c) self-report scores correlate with those of subordinates. Criterion-related validation is not demonstrated nor is there information on how the instrument relates to education.

The Hanson Silver Management Style Inventory and the Profiles from Rensis Likert Associates, Inc., may be of some value depending on the needs and interests of a test administrator. A self-report instrument, the Hanson Silver Management Style Inventory measures decision-making preferences based on Jung's theory of psychological types. One of the few group measurement

instruments, the Profiles for Rensis Likert Associates, Inc., provides the following files: (a) Conflict Characteristics, (b) Group Member Behavior, (c) Group Problem Solving, (d) Leadership Behavior, (e) Organized Characteristics, and (f) Own Behavior.

Assessment Center Methodology

Assessment centers are used for the selection and development of administrators and managers at entry- and mid-levels. Because identification and development of successful managers is of critical importance to any agency, business, or institution, this method was developed to increase the probability of successful selection.

Origin

After World War I, the German military forces were dismantled and the officer corps was dispersed. In the 1920s and '30s, the Germans began to rebuild their army, air, and naval forces and sought better ways of selecting officers to command units. Although forbidden to rebuild the military, the Germans first used civil agencies, as the postal service, to develop selection processes for leadership positions. Psychologists played a dominant role in the German programs and relied heavily upon holistic observations. Although many of their efforts were crude, e.g., observations were not recorded objectively, the

Germans introduced new means, such as leaderless group activities, for identifying men with potential for command.

The success of the Germans was noted by a British military attache in Berlin and subsequently the British adopted the method and established War Office Selection Boards (WOSBs). Whereas officer selection had been limited to men who had attended elite British schools, the WOSBs opened up selection to a broader social class. The assessment center method was quickly adopted by other members of the United Kingdom which showed the universality of the method. The British made several contributions in refining leaderless group activities and in relying less upon psychologists. Features of the War Office Selection Boards included discussion problems, indoor and outdoor activities, short speeches, physical activities, and stressful situations. Another major contribution was that analyses of selection were conducted; one finding was a correlation of .60 with War Office Selection Board grades and training grades (Thornton & Byham, 1982).

In the United States, pride of first place in assessment belongs to the Office of Strategic Services (OSS). In 1943, Henry Murray, the former head of the Harvard Psychological Clinical, and others were asked to develop means of selecting individuals to become spies, saboteurs, and support personnel for the OSS. Murray had conducted extensive personality studies at

ERIC

Harvard, and almost overnight, he and his colleagues developed ingeneous means of selecting agents. In addition to paper-and-pencil tests, inventories, and essays, extensive information was obtained by having each candidate engage in interviews and situational exercises. Another element of realism was added by requiring each candidate to maintain an assumed identity during the multi-day assessment. Without job descriptions for spies or saboteurs and with few candidates with experience in such occupations, the assessors had to develop novel means of selection so that the persons selected had a reasonable chance of returning from an assignment. For example, in "The Brook," a group of candidates was given the task of moving objects across a small stream that flowed through the estate in Virginia where the assessments were conducted. Opportunities for observing leadership, stress tolerance, and interpersonal skills were amply afforded by such activities. Other exercises, as "Behind the Barn," were developed to assess similar dimensions.

After WWII, the assessment center method was continued in Britain through the establishment of the British Civil Service Selection Board which was modeled after the War Office Selection Boards (WOSB). In the United States, the Veterans Administration Clinical Psychology Studies used the method extensively. Donald MacKinnon, who had served on the OSS staff, headed to

California and was instrumental in establishing the Institute for Personality Assessment and Research (IPAR) at Cal-Berkeley. The staff of the Institute conducted many job analyses and linked these effectiveness criteria. The California Psychological Inventory (CPI) has its roots in research based at the Institute for Personality Assessment and Research.

Psychological Testing

Advantages. Great benefits can accrue to employers and employees alike if the right persons are selected for appropriate jobs. Employers benefit if their selection processes enable them to hire individuals who become productive, loyal workers. Likewise, candidates for positions benefit if selection processes fit them with their interests, skills, and abilities. Little benefit comes from mismatches. The greater the mismatch, the more likely employees will leave or be forced to leave; thus employers must repeat the selection process while their former employees must spend their time, efforts, resources, and psychic energies on finding new employment.

Can psychological tests contribute to the match of persons with the job? At the plus end of the scale are those who answer, "Yes."

> . . . perhaps the greatest accomplishment of psychology thus far has been the measurement of the individuality of man. We know that we can, through various psychological techniques

assess the individuality of persons and make rather good predictions about their future behavior. Perhaps the greatest recent success attesting to this is the selection of Peace Corps candidates who are now serving overseas. I do not know what the current "hit rate" is for the Peace Corps selection project but I do know that of the first 600 persons who went overseas only 6 had to be returned and that 2 of these returned because of a death in the family. I would say the assumptions that "mind" or behavior cannot be measured is entirely negated by the evidence we have at hand (Dunnette, 1964, p. 67).

Psychological testing benefits from the aura of objectivity accorded to testing generally and from research findings that document test results with an external criterion, as job success. The objectivity of testing compares favorably with the subjectivity of clinical and supervisory evaluations. Evaluations of job performance by supervisors, in particular, suffer from a paucity of research findings.

The use of tests in identifying potential managers can be attractive, especially to the uninformed. Care must be taken so that tests are used properly, relate to job success, measure essential managerial factors, and can be interpreted properly.

ERIC

Disadvantages. Obtaining a successful "hit rate" requires the overcoming of many obstacles. The design and validation of an effective testing program is very difficult as job requirements must be clarified, an appropriate test or testing program must be developed or selected and if something beyond content validity is necessary, performance criteria must be established between the job and test.

Provisions for administering, scoring, and interpreting test results must be set. Training must be given to individuals who will make selection or other career decisions from test results so that testing programs are viewed as useful, fair, and accurate. Personality measures and general mental ability tests may not be perceived as relevant to job requirements. Test takers should be led to understand how a test measures what it is supposed to measure. Consequently, test administrators and users have a responsibility to test takers to show the relevance of tests.

Factors related to measurement and test construction, administration, scoring, and interpretation are important considerations but societal and legal matters must also receive attention. Increasingly, tests are expected to comply with legal considerations. The Equal Employment Opportunity Commission (EEOC) has guidelines regarding test validation so that potential

ERIC

racial discrimination or other forms of bias do not ensue from invalid testing.

Test development, selection, preparation of materials, administration, scoring, interpretation, research, and related factors all carry a cost. How the benefits from testing compare with the real and hidden costs is a subject that is difficult to answer. The search for an effective cost: benefit ratio propels many efforts to find more valid, better, cheaper, nondiscriminatory, and quicker tests.

Some tests, e.g., aptitude and mental ability, are objective, reliable, and readily standardized; however, interest and personality inventories are less so. "Personality tests can make valid contributions to the predictions of managerial effectiveness if they are carefully chosen, validated and interpreted by skilled psychologists. Past research evidence gives little indication of which tests are likely to be valid for which jobs" (Thornton & Byham, 1982, p. 72).

Testing of Management Potential

Advantages. A perceived advantage of assessment centers is that simulations evoke "samples" rather than "signs" of behavior. In the case of some paper-and-pencil tests, the only behavior that is required may be penciling in a response bubble on a multiple-choice answer sheet. While a test of general mental

ERIC

ability may indicate that a person is intelligent, i.e., has a high IQ, there is no guarantee that such a person will behave "intelligently." The brightest of persons fail to remember where the car is parked, to feed the parking meter, or to signal for a lane change. More complex demands produce more monumental lapses of behavior from "intelligent" people.

A driver's performance on a written examination may show knowledge (sign) of the rules of the road, but a trip behind the steering wheel lets an examiner observe actual skills in starting, pulling into traffic, maintaining proper distance and speed, parallel parking, turning corners into the proper lane, stopping, and showing courtesy to other drivers (sample). Likewise, situational tests let candidates exhibit their skills as potential administrators.

Applicants for a secretarial position who score highest on a paper-and-pencil vocabulary test (sign) will not necessarily make the fewest errors in preparing a business letter (sample). Because the complexity of managerial and administrative skills exceed simple tasks, candidates for administrative positions should be asked to perform tasks that elicit samples of behavior that are akin to those on the job. With its emphasis on job-related activities, the assessment center method appeals to participants and to hiring officials. The psychometric features

of assessment add to the method's acceptability. The use of multiple assessors, for example, helps to reduce error and bias.

Although the cost of assessment centers may appear to be high, they are cost effective in relation to performance with job-criteria (Cascio & Silbey, 1979). The utility of assessment centers is readily apparent when candidates have little or no opportunity to demonstrate managerial skills on their jobs or when supervisors have little basis for identifying potential administrators from among their subordinates.

The Graduate Record Examinations Board has recognized the limitations of traditional tests and is considering the use of the assessment center method to identify characteristics of successful students. The goal is "to work toward a more comprehensive description of what contributes to success in graduate school and to consider methods of substantiating this description" (Enright & Gitomer, 1989, p. 3).

Leaderless group discussions (LGDs). After the in-basket, leaderless group discussions are among the most widely used exercises in assessment centers (Thornton & Byham, 1982). Is there an administrative or managerial role that does not require participation in some group problem-solving activity?

Originally conceived by J. B. Rieffert, director of German military psychology (Thornton & Byham, 1982), LGDs have been used

by the War Office Service Boards (WOSBs), the Office of Strategic Services (OSS), and virtually every assessment center since. Leaderless group discussions can bring out samples of leadership behaviors, particularly if the problem or task is relevant to the participants or the job and can predict potential for leadership. Performance in LGDs can be affected by variation of participants in experience, rank, or verbal skills or, conversely, if participants are too much alike in leadership potential. Goal facilitation of the group, prominence, and sociability are three factors measured by leaderless group discussions (Cronbach, 1970).

Researchers have found good evidence of reliability and validity for leaderless group discussions. From the studies that Thornton and Byham (1982) have summarized, they show reliabilities from .64 to .71 for inter-rater rankings or ratings and validities as high as .66 for leaderless group discussions with job performance ratings.

Disadvantages. Several factors account for the disadvantages of assessment centers. A major difficulty is defining the job of an administrator or manager. What are the important tasks that a manager does? How can those tasks be sampled? How can subtle differences be accounted for, as in choosing principals or assistant principals to fill vacancies in several different schools in different neighborhoods? Because of the multitude and complexity of questions that must be addressed, assessment centers

are neither easy nor cheap to conceptualize, create, validate, and operate.

Because an assessment center typically lasts only a couple of days, how can behaviors that would appear over a long time be measured? Can preparing, administering, and evaluating a budget be squeezed into an hour-long activity? Must such a long-term cycle be incorporated into an assessment center? To what extent can an assessment center anticipate job demands from superordinates, subordinates, and organizational and environmental conditions? How many variables need to be taken into consideration? As an illustration, a boss's inclination to help and mentor a neophyte administrator can produce different on-the-job performance than another boss's philosophy of "sink or swim." To what extent can measures for a job be expected to correlate with performance on a job that likely has so many intervening variables, as Thorndike found in 10,000 Careers?

Reliability and validity remain a problem for assessment centers as they do for every exercise, test, or interview. Reliabilities and validities are often low to moderate and, at first glance, seem to be so low as to be meaningless; nevertheless, a detailed examination may reveal quite the opposite. Schmitt and Cohen (1990) found that the assessment center of the National Association of Secondary School Principals

(NASSP) has a criterion-related validity of .25. Of the participants who were given a placement recommendation of five (tops on the scale), 90% were given performance ratings on the job of five or higher on a six-point scale. Those who received a placement recommendation of two on the five-point scale were rated lower on job performance measures as only 48 percent received a rating of five or six. In brief, overall assessment ratings of "high fliers" were nearly twice as likely to result in high flying performance on-the-job in comparison to those whose overall assessment ratings were below average.

Reliability and validity of paper-and-pencil tests, interviews, and other measures should be compared with those for assessment centers. Often, criticism of the low reliability and validity of assessment centers is voiced without reference to data for other measures. The interview, for example, has relatively low validity.

> The inevitable conclusion derived from a number of investigations is that interview judgments, as they are usually made in the employment situation, are not closely related to independent measures of the characteristics judged. Nor are they closely related to measures of success on the job. In an overall sense the evidence regarding the validity

of the selection interview yields a distinctly disappointing picture (Miner & Miner, 1969, p. 277).

Yet correlations from selection interviews with subsequent promotion for a firm that had an "up or out" philosophy ranged from a -.12 for "character and habits" to .54 for "special comments" written at the conclusion of interviews (Miner & Miner, 1969). Some other correlations were -.10 for "technical competence" and .04 for "effectiveness with people." Predicting how people will behave on-the-job is not easy, particularly from an interview that may yield a few "signs" of behavior.

Leaderless group discussions. The assigned role (competitive) leaderless group discussion is criticized because few jobs require entry or lower-level administrators to be in competitive situations with peers. Another criticism is that behavior of participants can be affected unduly by the makeup of the group. In groups where a heady spirit of cooperation uniformly prevails, opportunity for expression of leadership skills varies from those groups whose participants range from passive to belligerent.

Other interactions of a group may also influence behaviors and mask or limit opportunities for display of leadership. An individual with exceptional leadership skills may make assessment of other participants in the group difficult; once again, the skill of the assessors is of vital importance. Excellent oral communication skills and forcefulness of a participant can combine

into vigorous actions that may unduly influence leadership ratings.

The Management Progress Study

In the mid 1950s, AT&T launched a landmark research project, the Management Progress Study (MPS), headed by Douglas W. Bray. Based on principles of the OSS assessment center, the MPS originally planned to assess beginning managers and follow their careers for the next eight years. The study has received much recognition because some participants have been reassessed twice, in years 8 and 20, and others have been followed into retirement. The Management Progress Study is still in operation as MPS:20 (Howard & Bray, 1988). The research was initially conceptualized as a study of adult development, and consequently, none of the information obtained from assessment was used for purposes of selection, promotion, or other career decisions. For the first 20 years, not even the participants were fed back information on how they had done lest the knowledge of self might contaminate the study.

In the Management Progress Study, Bray and Grant (1966) found that "The simulations--group problems and in-basket--show generally higher correlations (to staff prediction) than the paper-and-pencil devices. Among the latter, the mental ability test shows up, on the average, stronger than the personality

ERIC

questionnaire" (p. 15). Bray and Grant (1966) also reported interrater reliability of ratings at .75 and a validity of .60 with the overall assessment ratings (OAR) for the leaderless group discussion (LGD). In a later report, Bray (1973) indicated that interpersonal skills can be most effectively assessed through group exercises.

The results of the MPS became widely known among industrial psychologists. Most of the Fortune 500 companies and countless other agencies, businesses, governments, and military forces have adopted the assessment center method for selection or developmental purposes.

In the Management Progress Study, personality and motivational characteristics were derived from 37 scores from such instruments as the Edwards Personal Preference Schedule, the Guilford-Martin Inventory (GAMIN), the Sarnoff Survey of Attitudes Toward Life, and the Bass version of the California F Scale. In addition to interviews, the Thematic Apperception Test and two incomplete sentences tests (Rotter and Business Incomplete) were administered. From the 37 scores, six personality and motivation factors were produced by factor analysis. One scale--Leadership Motivation--was interpreted as follows:

High projective ratings on Leadership Role and Achievement/Advancement Motivation; low projective ratings on

ERIC

> Subordinate Role, Dependence, and Affiliation. High scorers on this factor have positive reactions to a leadership role and a desire for career accomplishments. They dislike a subordinate role or being dependent on others (Howard & Bray, 1988, p. 44).

After tracking the subjects for 20 years in the MPS:20 (Howard & Bray, 1988), several factors were found to be significantly related to success as measured mainly by promotion within AT&T. Ambition was the factor most closely related to career advancement ($r = .37$) while Leadership Motivation ($r = .10$), particularly for the noncollege men, and positive mental health--Self-Esteem ($r = .12$) and Positiveness ($r = .15$) were other characteristics of individuals who had climbed the organizational ladder.

Five of the seven dimension factors showed significant relationships with promotions after 20 years: (a) two motivational factors--Advancement Motivation ($r = .28$) and Work Involvement ($r = .15$) and (b) three ability factors--Intellectual Ability ($r = .25$), Interpersonal Skills ($r = .22$), and Administrative Skills ($r = .22$). Two personality dimension factors, Stability of Performance ($r = .12$) and Independence ($r = .00$), were not significantly related to promotion level. "Other variations in personality and attitudes,

including love of the company, were mostly irrelevant" (Howard & Bray, 1988, p. 77).

From the factor analysis of the data, promotion within AT&T was significantly related to motivation factors, intelligence, and interpersonal and administrative skills. Not many students of administrative theory would be surprised by such results. What may be surprising is that (a) these factors did appear from statistical analysis from the longitudinal study and that (b) the correlations were modest, with a range from .15 to .28.

Summary and Conclusions

Factors associated with personality and leadership have evoked keen interest in theorists and practitioners alike. The range of interest extends from that of theorists, who are primarily interested in abstract constructs, to that of practitioners, who want to apply knowledge of personality or leadership to critical problems of selection, promotion, or training of applicants to key positions within their organizations.

Principles of Measurement

Theorists and practitioners who want to measure complex concepts, such as human characteristics, have many obstacles to overcome. Identification, isolation, and measurement of a human quality, such as personality or leadership, are compounded by the difficulties of comparing two or more qualities in one person or

of comparing one or more qualities among individuals. Nevertheless, employing officials are interested in using personality and leadership factors in hiring, promoting, and training individuals to administrative and managerial positions. The increasing demands of a technological, globally interdependent society call for better selection processes so that effective leaders can be appointed to key positions. The research on effective schools, for example, contains ample documentation of the importance of the building principal in the education of students.

Advances in science have permitted physical scientists to measure objects from the size of atoms to galactic space, to chart movements as slow as that of the earth's crust to the speed of light, and to assess other forces in nature with varying degrees of precision. Variance in measurements in some cases, as in the movement of the earth's crust, are of less concern than others, as estimations of the date and degree of the next earthquake on the San Andreas Fault. Behavioral scientists have also established means to measure human attributes although their measurements may not be as exact as some made by natural scientists. Tolerance in the behavioral sciences for some errors is no less critical than predictions regarding the occurrence of earthquakes, volcanos, hurricanes, or tornados. Yet individuals place great confidence

ERIC

in scientific study, data analysis, tests and their results. Who can argue with the reading on the thermometer, the mechanic's engine analyzer, or the score from the Graduate Record Exam? Test scores are highly prized because they are, presumably, objective, precise, and so . . . so indisputably quantified. Although the speed of light and some physical laws are immutable, not all physical factors react in predictable ways; similarly, variables and qualities within an individual and among persons are not so constant. But the search goes on for identifying, classifying, and rating behaviors, traits, skills, and characteristics that are important for leaders to have.

Because instruments are less than perfect, measures of personality or leadership must show validity and reliability to be useful. Corporate executives of a food company that would have great variation of weights in its packaged cereals or other products would soon face the loss of customers, corporate credibility, profit, and their own jobs. Administrators who select employees by a process that would invariably put "square pegs in round holes" would soon face disgruntled workers, wary applicants, and dysfunction within the ranks.

Validity refers to the degree to which an instrument measures what it is supposed to measure. An instrument that purports to measure personality, for instance, must show evidence of construct

theory, that is, that some underlying principles account for performance on the instrument. A test may also need to demonstrate other forms of validity, such as criterion-related validity, to correlate with an outside measure as job performance. Leadership tests seldom have criterion-related measures so in selecting a test, preference may be given to a leadership test that has even a weak criterion-related measure to tests that have none.

A measurement tool is said to be reliable if it will show comparable results when administered under the same conditions. Reliability may be determined by several means, as administering the same test several days or weeks apart and subsequently computing the reliability coefficient.

Good tests have other characteristics. Major concepts or constructs must be supported by test results. The age, sex, race, number of subjects, level of education, and other demographic factors must be provided in the test manual so that test users can interpret results wisely. A test user must select a test with great care, administer the test as directed, and interpret scores with caution. Improper administration of a test can, for example, void the usefulness of a test.

Personality Tests

Everyone knows what personality is because everyone has one, but the concept of personality is so elusive that some may not even agree with that gross generalization. Many instruments designed to measure personality for large numbers of examinees are paper-and-pencil tests. Many of the well-known personality tests have been in print for several decades and are widely used. Each personality test has its own purpose and characteristics. Some were normed on "normals" or on populations that may not match well with an intended audience. Many have subscales that purport to measure discrete variables that may or may not fit with a test user's needs. Consequently, a test user must understand the nature of a personality test and interpret results with great care; for example, an inventory may have relatively low internal consistency, measure a few dimensions, or have limited data on norming for the targeted population.

Leadership Tests

Just as everyone knows what personality is, everyone knows what leadership is--until time comes to define and measure it. Leadership studies have included specific leadership skills, personality, physical characteristics, and other factors. What exactly constitutes leadership is not known as few, if any, traits appear to be necessities for effective leadership.

ERIC

Many theories have been postulated regarding leadership. Power and influence theories have focused upon the influence of leaders and the interactions between leaders and their respective followers. Behavioral theory is concerned with what leaders actually do. In the Ohio State University studies, for example, two key orientations--task and relationship--were the focus of study. Contingency theory focuses more upon external factors and includes the nature of the followers, the task, and the environment. Cultural and symbolic theories explore still other aspects of leadership. Research on leadership has considered psychological, sociological, and behavioral approaches; each approach has provided but a piece of the puzzle. One conclusion that has wide recognition is that no one leadership approach is best; nevertheless, not all approaches are equally good. Because leadership is apparently composed of a highly complex set of interrelationships, individuals looking for leadership are most likely to find bits and pieces rather than an holistic set.

The testing of leadership poses similar questions as personality. Data regarding purpose, constructs, scales, validity, reliability, norms, and other variables must be available to test users for proper selection, administration, scoring, and interpretation of results.

Assessment Center Methodology

Assessment centers are based upon the premise that behavioral dimensions that are job-related can be identified, observed, and rated by trained observers. Situational tests, like leaderless group discussions and group activities, are used to elicit leadership and other interpersonal behaviors. Assessment centers can elicit "samples" of behavior whereas paper-and-pencil tests generally provide "signs" of behaviors. For example, trained assessors can observe how participants interact, or display leadership, with each other in a group activity. Assessment centers pose particular problems also as behaviors may shift with a change in the composition of the group or with the nature of the group activity. Because assessment centers are more labor intensive, their costs are greater than paper-and-pencil tests. Assessment centers do have a long history and are a useful source of data about selected behaviors.

Conclusions

At present, the assessment center method is effective for direct observation of job-related skills. Paper-and-pencil tests and interviews have their particular strengths and offer economical means of providing data about examinees. Test users, however, must choose, administer, score, and interpret tests, whether they be paper-and-pencil or situational ones, with great

ERIC

care. To the uninitiated, a personality test may seem to be appropriate but actually may not measure what is thought to be measured. Neither is a paper-and-pencil test of leadership a good indicator of what a hiring official is looking for in this or that position. Nevertheless, some personality and leadership tests have lengthy histories of use and research behind them.

Too few studies have linked personality measures with leadership. Personality is believed to contribute to leadership success but the linkage needs additional research. Because personality and leadership have distinct theories behind each, comparisons of the two are compounded by measurement problems. Some aspects of personality, as dominance, appear to be important to leadership, but a combination of factors in one individual may produce leadership that is perceived to be more or less "equal" to that of another individual whose characteristics are dissimilar.

The potential of assessment centers in education is relatively new. The development of research on the relationship between personality and leadership of school administrators is an untapped but, likely, fruitful field. For the present, test users are advised, "Caveat emptor!" "Let the buyer beware," is sound advice to test users, hiring or promoting officials, staff developers, and others who are interested in the recruitment, selection, promotion, and development of school administrators.

ERIC

The use of many measures of assessment, including traditional means as letters of reference, transcripts, resumes, interviews, mental ability tests, and autobiographical sketches can contribute to assembling perceptions of individuals' strengths, skills, traits, and areas in need of improvement. Each instrument or method must be used with fidelity so the results are not compromised.

Tests must be selected carefully so that they match the intended purpose of the test user, are administered according to the instructions in the manual, are scored precisely, and interpreted wisely. If a test user is uncertain about a test, consultation should be obtained from an individual who has specific training in test administration, scoring, or interpretation. The right test, i.e., one that fits the purpose of the examiner, if administered properly, if scored with exactitude, and if interpreted correctly, will yield good results. The reverse is equally true. Individuals who want to ignore the utility of testing may, of course, rely on their own intuition and perceptions. Individuals relying on their hunches about the personality and leadership of applicants, however, should realize that they are taking great risks.

Administrators who are responsible for recruiting, screening, selecting, training, and promoting applicants for administrative

positions would do well to review the personnel procedures that they use to see how well their purposes are stated and the degree of validity and reliability they possess. Likewise, a review of interview techniques should be conducted to determine how well performance on-the-job is predicted from interviews. The perennial search for the "right personality" likewise needs better description, analysis, and validation. The selection of administrators might be improved by exerting more effort and resources in the design of the process and by using measures that are more appropriate for the task at hand.

References

Alker, H. A. (1978). [Review of the Minnesota Multiphasic Personality Inventory]. In O. K. Buros (Ed.), The eighth mental measurements yearbook (pp. 931-935). Highland Park, NJ: Gryphon Press.

Allport, G. W. (1966). Traits revisited. American Psychologist, 21(1), 1-10.

Anastasi, A. (1972). [Review of Personality Research Form]. In O. K. Buros (Ed.), The seventh mental measurements yearbook (pp. 297-298). Highland Park, NJ: Gryphon Press.

Baucom, D. H. (1985). [Review of the California Psychological Inventory]. In J. V. Mitchell (Ed.), The ninth mental measurements yearbook (pp. 250-252). Lincoln, NE: Buros Institute of Mental Measurements.

Barnhouse, R. T. (1984). Identity. Philadelphia: Westminster.

Bass, M. M. (1985). Leadership and performance beyond expectations. New York: The Free Press.

Bass, B. M. (1981). Stogdill's handbook of leadership: A survey of theory and research. New York: The Free Press.

Bennis, W. (1989a). On becoming a leader. Reading, MA: Addison-Wesley.

Bennis, W. (1989b). Why leaders can't lead. San Francisco: Jossey-Bass.

Bennis, W., & Nanus, B. (1985). Leaders. New York: Harper & Row.

Bensimon, E., Neuman, A., & Birnbaum, B. (1989). Making sense of administrative leadership: The "L" word in higher education. Prepared by ERIC in cooperation with ASHE (Association for the Study of Higher Education). Washington, DC: George Washington University.

Bernard, H. W. (1974). Personality/applying theory. Boston: Holbrook Press.

Birnbaum, R. (1989). How colleges work: The cybernetics of academic organization and leadership. San Francisco: Jossey-Bass.

Blake, R., & Mouton, J. (1964). The managerial grid. Houston: Gulf Publishing.

Bolman, L. G., & Deal, T. E. (1991, April). Reframing organizations: Artistry, choice and leadership. San Francisco: Jossey-Bass.

Bossert, S., Dwyer, B., Rowan, B., & Lee, G. (1982, Summer). The instructional management role of the principal. Educational Administration Quarterly, 18(3), 34-64.

Bray, D. W., & Grant, D. L. (1966). The assessment center in the measurement of potential for business management. Psychological Monographs, 80(17, Whole No. 625), 1-27.

Burns, J. M. (1978). Leadership. New York: Harper.

Cascio, W. F., & Silbey, V. (1979). Utility of the assessment center as a selection device. Journal of Applied Psychology, 64, 107-118.

Chapple, E. D., & Donald, G., Jr. (1946). A method for evaluating supervisory personnel. Harvard Business Review, 24, 201-203.

Covey, S. R. (1989). The seven habits of highly effective people. New York: Simon & Schuster.

Cronbach, L. J. (1970). Essentials of psychological testing. New York: Harper & Row.

Deal, T., & Kennedy, A. (1982). Corporate cultures: The rights and rituals of corporate life. Reading, MA: Addison-Wesley.

Dingle, H. (1960). Basic problems of measurement. Scientific American, 202(6), 189-192.

Dipboye, R. L. (1978). [Review of the Leader Behavior Description Questionnaire, Form 12]. In O. K. Buros (Ed.), The eighth mental measurements yearbook (pp. 1750-1751). Highland Park, NJ: Gryphon Press.

Duke, D. L. (1987). School leadership and instructional improvement. New York: Random House.

Dunnette, M. D. (1964). Critics of psychological tests: Basic assumption: How good? Psychology in the Schools, 1(4), 63-69.

Eberhardt, B. J. (1978). [Review of Situational Leadership]. In O. K. Buros, (Ed.), The eighth mental measurements yearbook, (pp. 1384-1385). Highland Park, NJ: Gryphon Press.

Enright, M. K., & Gitomer, D. (1989). Toward a description of successful graduate students. GRE Board Research Report No. 85-17R. ETS Research Report 89-9. Princeton, NJ: Educational Testing Service.

Erikson, E. H. (1978). Adulthood. New York: Norton.

Fiedler, F. E. (1967). A theory of leadership effectiveness. New York: McGraw-Hill.

Fiedler, F. E. (1971). Validation and extension of the contingency model of leadership effectiveness: A review of the empirical findings. Psychological Science Quarterly, 76(2), 128-148.

Fiedler, F. E., & Garcia, J. E. (1987). New approaches to effective leadership. New York: Wiley.

Fisher, R., & Brown, S. (1988). Getting together: Building relationships as we negotiate. London: Penguin.

Foote, N., & Cottrell, L., Jr. (1955). Identity and interpersonal competence: A new direction in family research. Chicago: University Press.

Gardner, J. W. (1990). On leadership. New York: The Free Press.

Gibb, C. (1968). Leadership. In G. Lindzey & E. Aronson (Eds.), The Handbook of Social Psychology, Vol. 4. Reading, MA: Addison-Wesley.

Gough, H. G. (1968). An interpreter's syllabus for the California Psychological Inventory. In P. McReynolds (Ed)., Advances in Psychological Assessment. Palo Alto, CA: Science and Behavioral Books.

Groth-Marnat, G. (1990). Handbook of psychological assessment. New York: Wiley.

Henry, W. E. (1949). The business executive: The psychodynamics of a social role. American Journal of Sociology, 54, 286-291.

Hersey, P., & Blanchard, K. (1977). Management of organizational behavior (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hogan, R. (1978). [Review of Personality Research Form]. In O. K. Buros (Ed.), The eighth mental measurements yearbook (pp. 1007-1008). Highland Park, NJ: Gryphon Press.

House, R. J. (1971). A path-goal theory of leader effectiveness. Administrative Science Quarterly, 16, 321-338.

Howard, A., & Bray, D. W. (1988). Managerial lives in transition. New York: The Guilford Press.

Howell, J., Dorfman, P., & Kerr, S. (1986). Moderator variables in leadership research. Academy of Management Review, 11(1), 88-102.

Hoy, W. K., & Miskel, C. G. (1987). Educational administration: Theory, research, and practice (3rd ed.). New York: Random House.

Hughes, J. (1989). [Review of the Educational Administrator Effectiveness Profile]. In J. C. Conoley & J. J. Kramer (Eds.), The tenth mental measurements yearbook (pp. 280-282). Lincoln, NE: Buros Institute of Mental Measurements.

Kelly, E. (1972). [Review of Personality Research Form]. In O. K. Buros (Ed.), The seventh mental measurements yearbook, 298-301. Highland Park, NJ: Gryphon Press.

Kerlinger, F. N. (1979). Behavioral research: A conceptual approach. New York: Holt, Rinehart, and Winston.

Kerr, S., & Jermier, J. M. (1978). Substitutes for leadership: Their meaning and measurement. Organizational Behavior and Human Performance, 22, 375-403.

King, G. D. (1978). [Review of California Psychological Inventory]. In O. K. Buros (Ed.), The eighth mental measurements yearbook (pp. 935-938). Highland Park, NJ: Gryphon Press.

Kotter, J. P. (1982). The general managers. New York: The Free Press.

Kotter, J. P. (1988). The leadership factor. New York: The Free Press.

Kouzes, J. M., & Posner, B. Z. (1990). The leadership challenge. San Francisco: Jossey-Bass.

Lawrence, A., & Jaffe, D. (Eds.). (1977). Reading in adult psychology: Contemporary perspectives. New York: Harper & Row.

Lipham, J. M., & Hoeh, J. A., Jr. (1974). The principalship: Foundations and functions. New York: Harper & Row.

Lipham, J., Rankin, R., & Hoeh, J., Jr. (1985). The principalship: Concepts, competencies, and cases. New York: Longman.

McGregor, D. (1960). The human side of enterprise. New York: McGraw-Hill.

McKee, M. G. (1972). [Review of Edwards Personal Preference Schedule]. In O. K. Buros (Ed.), The seventh mental measurements yearbook (pp. 149-151). Highland Park: Gryphon Press.

Menninger, K. A. (1953). The human mind (3rd ed.). New York: Alfred A. Knopf.

Miller, L. M. (1984). American spirit: Visions of a new corporate culture. New York: William Morrow.

Miner, J. B., & Miner, M. G. (1969). Personnel and industrial relations: A managerial approach (2nd ed.). New York: Macmillian.

ERIC

Mintzberg, H. (1973). The nature of managerial work. New York: Harper & Row.

Norman, W. T. (1972). [Review of Edwards Personality Inventory]. In O. K. Buros (Ed.), The seventh mental measurements yearbook, (pp. 153-155). Highland Park, NJ: Gryphon Press.

Pfeffer, J. (1977). The ambiguity of leadership. Academy of Management Review, 2, 104-112.

Piatrowski, C., & Keller, J. W. (1984). Psychodiagnostic testing in APA-approved clinical psychology programs. Professional Psychology: Research and Practice, 15(3), 450-456.

Schmitt, N., & Cohen, S. A. (1990). Criterion-related validity of the National Association of Secondary School Principals' Assessment Center. Reston, VA: National Association of Secondary School Principals.

Schroeder, J. E. (1986). Videodisc Interpersonal Skills Training and Assessment (VISTA) overview and findings (Vol. 1, Final Report). Alexandria VA: Training Research Lab., U.S. Army Research Institute for the Behavioral and Social Sciences.

Sergiovanni, T. J. (1990). Adding value to leadership get extraordinary results. Educational Leadership, 47(8), 23-27.

Shibles, M. R. (1988). School leadership preparation: A preface for action. Washington, DC: American Association of Colleges for Teacher Education.

Smith, P., & Peterson, M. (1988). Leadership, organizations and culture: An event management model. London: Sage.

Smith, R., & Monahan, W. (1990-91). Leadership and contextualism: Principles of military leadership applied to the administration of education. National Forum of Educational Administration and Supervision Journal, 7(3), 163-176.

Stogdill, R. M., & Coons, A. E. (1957). Leader behavior: Its description and measurement. Columbus, OH: Bureau of Business Research, Ohio State University.

Tellegen, A. (1978). [Review of Eysenck Personality Inventory]. In O. K. Buros (Ed.), The eighth mental measurements yearbook (pp. 802-804). Highland Park, NJ: Gryphon Press.

Thorndike, R. L. (1963). The prediction of vocational success. Vocational Guidance Quarterly, 11, 179-187.

Thorndike, R. L. (1964). Educational decisions and human assessment. Teachers College Record, 66(2), 103-112.

Truman, H. S (1989). Where the buck stops. New York: Warner Books.

Vroom, V., & Yetton, P. (1973). Leadership and decision making. Pittsburgh: University of Pittsburgh Press.

Wendel, F. C. (1985). An introduction to assessment center methodology. Lincoln, NE: Department of Educational Administration, University of Nebraska.

Wendel, F. C., & Breed, R. D. (1988). Improving the selection of principals: An analysis of the approaches. _NASSP Bulletin, 72_, 35-38.

Westley, F., & Mintzberg, H. (1989). Visionary leadership and strategic management. _Strategic Management Journal, 10_, 17-32.

Wiggins, J. (1989). [Review of the Myers-Briggs Type Indicator]. In J. C. Conoley & J. J. Kramer (Eds.), _The tenth mental measurements yearbook_ (pp. 537-538). Lincoln, NE: Buros Institute of Mental Measurements.

Wilkins, A. L. (1989). _Developing corporate character_. San Francisco: Jossey-Bass.

Wrightsman, L. S. (1988). _Personality development in adulthood_. Beverly Hills, CA: Sage.

Yukl, G. A. (1971). Toward a behavioral theory of leadership. _Organizational Behavior and Human Performance, 6_, 414-440.

Yukl, G. A. (1981). _Leadership in organizations_. Englewood Cliffs, NJ: Prentice-Hall.